A conserved gram-positive signalling gene cluster: Phylogeny and function.

Jones, Greg

A Conserved Gram – Positive Signalling Gene Cluster: Phylogeny and Function.

A thesis submitted in candidature for the degree of PhD

Greg Jones

Swansea University
School of Medicine
www.medicine.swansea.ac.uk

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Acknowledgements

So many people have helped and supported me through my studies that it is difficult to know where to start. First of all, my lovely wife, Katie, has been a constant source of love and invaluable support. I couldn't have done this without her. Thanks are also due to Paul who has been a splendid supervisor due to his laid – back and knowledgeable style. The day – to day supervision, experience, support and muttered Cuban expletives provided by Ricardo were also essential to the completion of this thesis. Well, maybe not all of the swearing! Sue and Meirwyn provided a large amount of help as well for which I am deeply grateful. My fellow lab – rats deserve a mention just for putting up with me for so long, as well as all the help and friendship they gave me – so thanks Lorena, Sandra, Amy, Bazza, Simon and everyone else who helped to make my time in Swansea so enjoyable. Last, but by no means least, my family and family – in – law have been fantastic.

Summary

All sequenced actinobacteria to date contain a conserved gene cluster found close to the origin of replication of the chromosome. This gene cluster contains a number of genes devoted to both signalling on serine and threonine residues a serine/threonine protein kinase (STPK), two 'forkhead associated'(FHA) domains, and a ser/thr phosphatase as well as genes devoted to cell wall synthesis. Bioinformatic analysis reveals variations of the cluster conserved in firmicute genomes, implying a conservation of function. The key gene in this cluster is the STPK, as it is a transmembrane protein whose extracellular portion consists of four 'PASTA' domains which seem likely to be sensing peptidoglycan monomers associated with cell elongation. Mutation of the genes associated with signalling in *Streptomyces coelicolor* causes the early onset of secondary metabolism and aerial development, implying a link between the state of the cell wall and the control of metabolic flux. The supplementation of growth media with metabolic intermediates reveals a more direct link between FHA domains and metabolic flux, with excess citrate causing a particular increase in secondary metabolism. Further proteomic analysis confirms the involvement of proteins dedicated to metabolic functions, including the exclusive presence of phosphorylated citrate synthase in the FHA domain mutants. The emerging picture is one of a pleiotropic network of proteins, using serine / threonine phosphorylation as a signalling device, controlling metabolic flux in response to a variety of environmental signals.

Contents

1 Introduction

1.1 Why study Streptomyces?

According to the latest World Health Organisation (WHO) statistics, tuberculosis (caused by *Mycobacterium tuberculosis*) kills nearly two million people a year. In combination with the HIV/AIDS epidemic sweeping parts of Africa and Asia it becomes even more deadly as immunocompromised people succumb to it more easily. Poor medical facilities and the emergence of 'Extemely drug resistant' variants have contributed to the global pandemic (http://www.who.int). This puts studying *M. tuberculosis* at the forefront of the medical research agenda, and this research will also help design strategies against other disease causing actinobacteria such as *Mycobacterium leprae, Mycobacterium bovis, Tropheryma whipplei* and one of the causative agents of acne *Propionibacterium acnes*. However, studying the mycobacteria can prove problematic – they are slow growing, gene knockouts are often lethal and practical considerations arising from their pathogenicity can hamper research efforts. Luckily, the non – pathogenic *Streptomyces* share a large amount of genetic similarity with *Mycobacteria,* and the high degree of redundancy within the large genome can allow gene knockouts to be studied both biochemically and phenotypically. In addition to aiding research into TB, the *Streptomyces* are the source of most of the antibiotics used by humans (Watve et al. 2001), as well as being the most likely vehicles for newly synthesised compounds created to combat the spread of antibiotic resistant organisms. A thorough understanding of the genetics and physiology of the *Streptomyces* is thus essential for improving the health of millions of human beings.

1.2 Taxonomy of Streptomyces

Streptomyces coelicolor is a Gram positive filamentous bacterium that is found within the order *Actinomyces*. This order is very diverse, however the phylogenetic relationships between species has remained relatively constant as new molecular techniques were introduced, unlike many other areas of bacterial phylogeny (Anderson and Wellington 2001). The diversity of the Actinomycetes is reflected in their phenotype, habitat and molecular biology. For example, the genomic G+C content of *Tropheryma whipplei* (a human pathogen) is 46% (Bentley et al. 2003; Raoult et al. 2003), whereas that for *S. coelicolor* (found in soil) is 72% (Bentley et al. 2002). The genus *Streptomyces* is the source of the majority of all bioactive compounds produced by the Actinomycetes (over 9000 have been isolated), and as a consequence they are also the source of the majority of antibiotics.

1.3 Life cycle of S. coelicolor

S. coelicolor has a complex life cycle, involving three distinct morphological stages. First, a unigenomic spore germinates in solid media and the filaments produced branch and form a mycelial structure. Secondly, aerial hyphae are formed as the substrate mycelium is degraded to provide nutrients and energy for this process. Thirdly, the aerial hyphae undergo septation to form unigenomic spores which can then germinate once again.

Each stage of the developmental cycle is controlled by different genes, and the production of the various metabolites is closely linked to the developmental state of the colony. The molecular basis for this is covered in more detail in a later section.

1.4 Genome structure

The genomes of all *Streptomyces* examined so far are linear, and have covalently attached proteins at the telomeres. The termini consist of inverted repeats that vary in length from 24kb in *Streptomyces griseus* to 550kb in *Streptomyces rimosus* (Volff and Altenbuchner 1998). The two publicly available *Streptomyces* genomes from *S. avermitilis* and *S. coelicolor* show that both these organisms share a number of genomic features that are likely to be common to all members of the *Streptomyces*. The general arrangement of genes shows concentration of essential functions such as metabolic enzymes and cell division machinery in the central part of the chromosome. Secondary metabolism genes and other less vital functions are more likely to be found in the chromosome 'arms', and these regions are also more likely to be subject to rearrangements such as transposon insertion or recombination. The two genomes share 69% of their predicted ORFs, and both code for a large number of sigma factors (60 in *S. avermitilis,* 65 in *S. coelicolor*). The core region has more similarity to circular actinomycete chromosomes, and thus it seems likely that this region is the more ancient part of the *Streptomyces* genome. The first sequenced strain of *S. coelicolor* has a genome size of 8.7 megabases, and codes for 7825 ORFs (Bentley et al. 2002; Ikeda et al. 2003). Later examination of other strains, including M145, revealed terminal inverted repeats of 1Mb thus increasing the genome size to 9.7Mb (Weaver et al. 2004). To underline the diversity of function encoded, the number of ORFs is greater than that for *Saccaromyces cerevisiae* (6294 ORFs) and approximately a third of the number of genes thought to be necessary for a human being! The large number of biologically active secondary metabolites has evolved to discourage competition for scarce resources, and those resources are then acquired via a vast arsenal of extracellular enzymes targeted at relatively chemically inert

substrates such as chitin and cellulose. There must also be responses to rapid osmotic, temperature and oxidative changes in the immediate environment (for example a thunderstorm on a hot day would suddenly flood the soil and cause local anoxic conditions as well as the rapid change in osmotic pressure). The linearity of the chromosome is thought to be due to recombination with a linear plasmid (Chen et al. 2002).

S. coelicolor A3(2) plays host to three plasmids, SCP1, SCP2 and SLP1. SLP1 is usually found integrated into the chromosome, as it contains phage – like factors (*int* and *xis*) allowing it to integrate in a non – mutagenic, site specific manner with conserved tRNA genes. The plasmid can transfer itself autonomously to a strain lacking the plasmid, and it can replicate stably in a circular configuration if the *att* or *int* genes are disrupted. SCP2 (30kb) is a circular non – integrating plasmid that probably replicates from a bi – directional origin of replication. SCP1 (350kb) is a linear non – integrating plasmid that encodes synthesis of and resistance to methylenomycin (T. Kieser 2000). Some laboratory strains are lacking in one or more of these elements, for example strains M145 and M600 lack SCP1 and SCP2.

1.5 Molecular basis of differentiation

1.5.1 Re gulation of hyphal growth

The classic model of streptomycete growth starts with a unigenomic spore that germinates to form a germ tube which then grows by tip extension and branching to form a dense substrate mycelium. Aerial hyphae are then erected and these go on to form spores. As might be guessed this process involves a wide range of molecules which sense the environment, regulate gene expression and supply the physical

structures to effect the changes. Hyphal growth is propagated via tip extension, and current knowledge indicates that protein DivIVA is a major part of the apical extension complex (Flardh 2003), as it localises in high concentration at the hyphal tips. Overexpression of this protein resulted in swollen apical tips and hyperbranching. Partial depletion of DivIVA resulted in atypical 'curly' hyphae and branching in the apical cell of substrate hyphae, and null mutants are lethal. There is evidence for DivIVA localising immediately after germination, and so it appears that the protein is integral to both tip extension and tip localisation (the swelling in strains overexpressing the protein are most likely due to multiple extension sites being present). Cell wall assembly occurs predominantly at the tip, and thus penicillin binding proteins (PBPs) and other cell wall assembly proteins must also localise there. Evidence for extension in areas other than the apical tip is sparse, however *S, coelicolor* and *S. avermitilis* contain many PBP genes, the protein products of which are located distally from cell poles in *Bacillus subtilis* (Scheffers, Jones, and Errington 2004) and these may well contribute at least to cell wall maintenance away from the tip if not to distal wall extension.

1.5.2 Re gulation of aerial development

If descriptions of hyphal growth seem to be lacking, it is because it is generally described in terms of repression of aerial development and secondary metabolism. The genes most associated with early stages of aerial development are the 'bald' (*bld*) genes, and the genes associated with later stages are the 'white' (*whi*) genes (named for the phenotypes observed when their representatives are mutagenised). These genes are involved in a regulatory cascade that responds to a wide variety of stimuli, and through this the onset of spore formation can be closely tied to a particular set of environmental parameters. The *bld* genes regulate the production of SapB (Claessen

et al. 2006), a secreted lantibiotic – like peptide that acts as a surfactant and lowers the surface tension at the interface between substrate hyphae and air. This evidence has been gleaned in part from extracellular complementation studies in which it has been shown that different *bld* mutants grown near to each other can rescue the formation of aerial hyphae. This has led to the cascade theory (Chater 2001). Present evidence shows the *bld* cascade having the following hierarchy:

b*ldJ* < *citA* < *bldK* – *bldL* < *bldA* – *bldH* < *bldG* < *bldC* < *bldD* – *bldM*

with mutants of genes to the right able to complement mutants of those on the left, however not *vice versa.* Production of SapB is specified by the *ram* (for 'rapid aerial morphogenesis') cluster. This cluster is under the control of *ramR*, overexpression of which bypasses the *bld* phenotype, however in the wild type *ramR* is under the control of the *bld* cascade (Nguyen et al. 2002).

Following SapB's lowering of the surface tension, the hyphae become coated with a layer of hydrophobic protein fibrils called the rodlet layer. These fibrils are made from individual monomers known as 'chaplins' which are also strong surfactants. The chaplins are arranged into fibrils by proteins known as rodlins, and these fibrils ensure that the aerial hyphae and spores remain in the air rather than collapsing back into the substrate (Claessen et al. 2004). The genes encoding both chaplins and rodlins (*chp* and *rdl* genes respectively) are under the control of the *bld* cascade, however there does seem to be an extra regulatory checkpoint dubbed the 'Sky' pathway. This pathway appears takes over regulation of aerial hyphae after induction of *ram* genes, *chpE* and *chpH* by the *bld* cascade, and this is supported by the complementation of a *ramS* mutant if an extracellular surfactant is added (Claessen et al. 2006).

Following the erection of aerial hyphae the next stage of development is the formation of spores, and this stage is primarily under the control of the *whi* genes, coupled with cell division genes that are universally conserved in bacteria. The control of sporulation is initiated by the sigma factor encoded by *whiG*, which is repressed by *bldD* (Elliot et al. 2001). Briefly, the 'early' *whi* genes (*whiA*, *whiB*, *whiG*, *whiH*, *whiI*, and *whiJ*) control the formation of septa in the aerial hyphae (Ryding et al. 1998; Flardh, Findlay, and Chater 1999). Other *whi* gene products are responsible for the unique features of the spores (Ryding et al. 1999).

1.5.3 Re gulation of secondary metabolism

Secondary metabolism is difficult to define, as a metabolic pathway that may seem superfluous in the laboratory may be essential in the organism's natural habitat. In studies of *S. coelicolor*, the term 'secondary metabolism' is often taken to be synonymous with 'antibiotic production', which is the approach I will take here. *S. coelicolor* produces four antibiotics, actinorhodin (act), undecylprodigiosin (red), calcium dependent antibiotic (cda), and methylenomycin (see section 1.3). Secondary metabolism coincides with aerial growth, and the most plausible theory for this coupling of the two processes is that antibiotics will suppress competition from other microbes as the substrate hyphae lyse and release their nutrients for use in the aerial hyphae. As colony development and secondary metabolism are temporally linked, it is no surprise that there is a regulatory link as well. The *bld* genes are required for antibiotic production, as *bld* mutants do not produce antibiotic (T. Kieser 2000), and if polyketide production in general is considered as secondary metabolism, mutation of *whiE* prevents the grey polyketide spore pigment from being formed (Kelemen et al. 1998). In addition to developmental cues, secondary metabolism can be controlled by environmental signals, or by extracellular signalling molecules such as γ-

butyrolactone molecules (Chater and Horinouchi 2003; Takano 2006). These molecules are important extracellular signalling molecules, and their function was first investigated with respect to 'A-factor' in *Streptomyces griseus*. In this organism, the role of A-factor is to initiate aerial development and secondary metabolism after its concentration has passed a critical level in the substrate mycelium. This is in contrast to the situation in *S. coelicolor*, where γ-butyrolactones have been found only to stimulate secondary metabolism (Ohnishi et al. 2005).

γ-butyrolactones are not the only signalling molecules used in the regulation of secondary metabolism. For example, it has been shown that accumulation of the multiply phosphorylated guanosine nucleotide ppGpp is responsible for activation of actinorhodin production (Hesketh, Sun, and Bibb 2001). This molecule is produced in *E. coli* from ATP and GTP by a ribosome bound ppGpp synthetase enzyme, RelA and a bi – functional ppGpp hydrolase, SpoT. The effect of its accumulation in *E. coli* is to change the pattern of transcription in times of stress by binding to RNA polymerase (Magnusson, Farewell, and Nystrom 2005). The advantage of this system is that only a single molecule is needed to activate a wide range of stress responses. In *S. coelicolor*, ppGpp accumulation is conditionally responsible for starting the transcription of actinorhodin and undecylprodigiosin synthesis genes (Hesketh, Sun, and Bibb 2001; Sun, Hesketh, and Bibb 2001). The synthesis of ppGpp is under the control of the homologues of RelA and SpoT found in *S. coelicolor*. ppGpp levels can be influenced by other genes, for example *eshA* can also modulate ppGpp levels, and its mutation leads to loss of antibiotic production due to lowered ppGpp levels. Production can be restored through overexpression of *relA* (Saito et al. 2006).

The signals, sensors and effectors of morphological differentiation in *S. coelicolor* have been the subject of intense study over the last decades, even so we are still discovering new signalling and regulatory processes. The end result of the various cascades is to influence development, and this is often seen as changes in cell wall arrangement (e.g. formation of septa in aerial hyphae). The changes in cell wall structure are underpinned by a highly conserved set of general biochemical processes, which in turn are under the control of some highly conserved genes.

1.6 Cell wall structure

1.6.1 T he dcw cluster

The following few sections contain an overview of the essentials of bacterial cell division, however before getting into the details, it is worth noting the genomic arrangement of the vast majority of the cell division genes. The *mur* genes are clustered with *fts* genes and various *pbp* genes. This cluster is called the division cell wall (dcw) cluster. Interestingly, rod – shaped cells have a much more conserved cluster arrangement than cocci, implying that the order of genes is important in terms of cell shape (Massidda et al. 1998; Mingorance, Tamames, and Vicente 2004). Transcription within it is complex, with many ORFs overlapping, with some genes (e.g. *ftsZ*) being transcribed from more than one promoter (Vicente, Chater, and de Lorenzo 1999).

1.6.2 A ssembly of peptidoglycan

Peptidoglycan is ubiquitous in nature, as the vast majority of eubacteria are sheathed in this macromolecule. Its main function is to withstand the high internal osmotic pressure that otherwise would rupture the cytoplasmic membrane. It is no coincidence

that organisms without peptidoglycan cell walls (e.g. *Mycoplasma*) are found in osmotically favourable environments. The other functions of peptidoglycan are to define cell shape and as an integral part of the cell division / growth processes.

The first stage in peptidoglycan synthesis is the formation of glycan chains from N-acetyl muramic acid (NAM) and N-Acetyl glucosamine (NAG). UDP-NAM monomers have successive amino acids added (this intermediate is known as LipidI), then a UDP-NAG monomer is added to them to form a NAG-(NAM-Pentapeptide) monomer (LipidII). Lipids I and II possess a lipid like 'tail' ($C_{55}H_{89}$) which is used to anchor the molecules in the cell membrane during transpeptidation and transglycosylation. This process involves the *mur* genes *A – F* followed by *mraY*, and then *murG*. The disaccharide peptide monomers are then joined by glycosyltransferases, to form a long macromolecule after being exported out of the cytoplasm across the membrane (van Heijenoort 2001). For a strong wall to be formed, it is imperative that the NAM-NAG backbone molecules be joined to each other, and this is where the pentapeptide is used. Transpeptidase enzymes join pentapeptides using the terminal D-Ala and adjacent di-aminopimelic acid from different backbone molecules (see figure 1.1) to form a web of peptidoglycan that envelopes the bacterial cell (Holtje 1998).

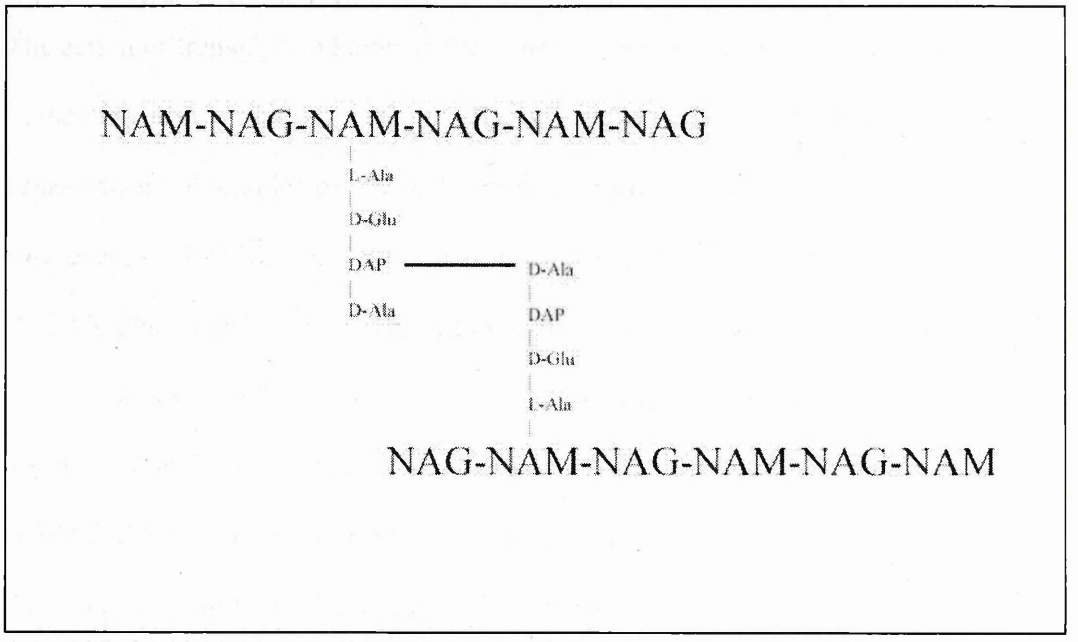


Figure 1.1: schematic diagram of linkage between cell wall strands, showing one linkage. DAP = diaminopimelic acid. The links between NAM and NAG subunits are β 1-4 links. The link shown between D-Ala and DAP can contain up to 5 amino acids, dependent on species.

The action of transglycosylation and transpeptidation is performed by proteins known as the penicillin binding proteins (PBP). Their name derives from the observation that beta – lactam antibiotics bind to these proteins preferentially (Spratt 1975). There are two classes of PBP: high molecular weight (HMW) and low molecular weight (LMW). The HMW PBPs are bifunctional molecules that can be split into two further groups. Group A PBPs contain an N-terminal transglycosylase and a C-terminal transpeptidase. These proteins are homologous to PBP1a and PBP1b in *E. coli*. Group B HMW PBPs retain the C-terminal transpeptidase activity, however their N-terminal domains have not been characterised. There is speculation that it is a morphological determinant, possibly interacting with other cell division proteins. PBP2 and PBP3 are *E. coli* representatives of this class (Scheffers and Pinho 2005). LMW PBPs seem less important than the HMW PBP1 class (Young 2001). The LMW PBPs are monofunctional enzymes that have endopeptidase, carboxypeptidase and beta lactamase activity. Their role in cell wall metabolism is unclear, as a form of peptidoglycan can be synthesised in *E. coli* when the majority of them are mutated (Denome et al. 1999).

1.6.3 Str ucture of peptidoglycan – horizontal or vertical?

The traditional view of peptidoglycan orientation has the NAG-NAM backbone lying parallel to the cell membrane, with the peptide bonds providing flexible cross links that allow expansion as the cell enlarges (Holtje 1998). This classical model of peptidoglycan has the backbone lying parallel to the cell membrane, with every fourth peptide bound in the same plane. This model proposes a '3 for 1' method of cell wall synthesis, with a multienzyme complex of transpeptidases and transglycosylases degrading a single strand of peptidoglycan and replacing it with three new strands as

it circles the cell using the strand to be degraded as a rail to follow. Recently, a new model has been proposed by Dmitriev and co workers, named the 'Scaffold' model (Dmitriev et al. 2003). This model has the NAG-NAM backbone orientated perpendicularly to the cytoplasmic membrane, and the peptide bonds are arranged in a hexagonal manner to form pores. Recent crystallographic evidence for the structure of Gram – positive murein has supported this idea (Meroueh et al. 2006), and the authors also note the near – perfect fit of a TolC protein into one of the hexagonal pores. Each model has its supporters and detractors (Vollmer and Holtje 2004; Young 2006), and currently the general assumption of a horizontal rather than perpendicular arrangement of glycan strands is prevalent.

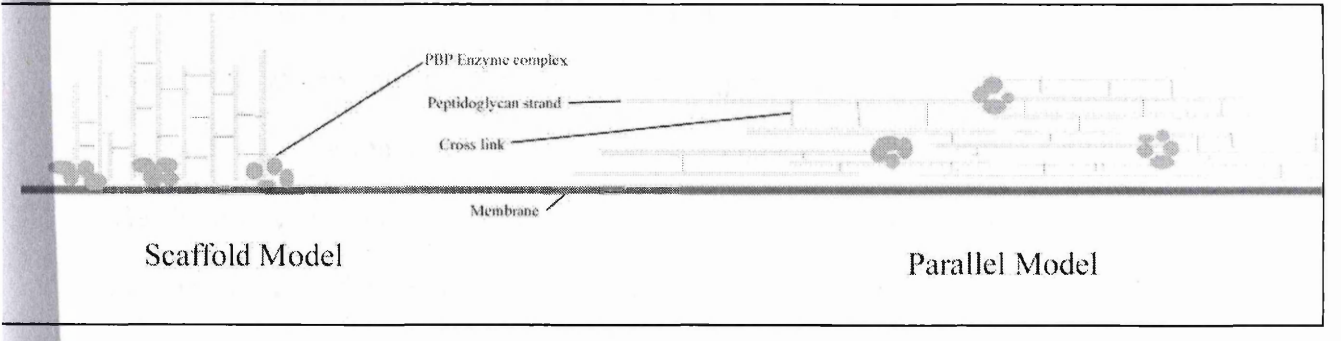


Figure 1.2: Two models of peptidoglycan assembly. The scaffold model relies on membrane associated enzyme complexes assembling strands at the point of precursor secretion. The parallel model has enzyme complexes assembling peptidoglycan as discreet units within the cell wall.

1.7 *Streptomyces* cell division

1.7.1 *fts* genes and cell division

ftsZ is one of the most highly structurally conserved and best studied genes in bacteria. It was first characterised in 1980 (Lutkenhaus, Wolf-Watz, and Donachie 1980), and it was found that *E. coli* mutants of this gene have a filamentous, conditionally thermosensitive phenotype (hence *fts*). This matched the phenotype observed for other mutations of cell division genes seen before. A decade later, it was established that *ftsZ* was essential for cell division and septum formation in *E. coli* (Dai and Lutkenhaus 1991). The study of *ftsZ* and its related genes has led to great advances in understanding the process of bacterial cell division, and has gone some way to providing a conceptual link between peptidoglycan synthesis and cell division. The following is a brief summary of this process, however the reader is strongly advised to read the review of Goehring and Beckwith and references therein (Goehring and Beckwith 2005) for a comprehensive overview.

S. coelicolor does not require *ftsZ* for growth, as an *ftsZ* mutant was able to grow and produce aerial hyphae. This showed that, in this organism, *ftsZ* is only employed in the formation of spores and the few vegetative septa formed in substrate hyphae (McCormick et al. 1994). It also seems that FtsZ is actively inhibited by proteins such as CrgA in vegetative hyphae (Del Sol et al. 2003; Del Sol et al. 2006). The FtsZ protein is a GTPase that forms a ring structure at the site of septation. Its structure is similar to that of eukaryotic tubulin, which is used in separation of chromatids during eukaryotic cell division. The Z - ring is in fact believed to be a tight spiral rather than

a closed hoop, and this structure is also highly dynamic with a high rate of flux between FtsZ in the ring and free proteins (Chen and Erickson 2005). The Z - ring also provides the scaffold on which the rest of the cell division machinery can be assembled (Goehring and Beckwith 2005).

Before the polymerisation of FtsZ units into the Z – ring, the site of septum placement must be determined. *E. coli* and *B. subtilis* use a conserved system to place the septum at midpoint, the *min* system. MinC and MinD inhibit FtsZ polymerisation. In *E. coli,* MinC and MinD oscillate from one cell pole to the other, so FtsZ localises to the point at which fewest MinCD molecules are located – i.e. the midpoint of the cell (Fu et al. 2001). *B. subtilis* has a different system, using DivIVA to anchor the MinCD proteins to the cell poles (Anderson, Gueiros-Filho, and Erickson 2004). Searches of the *S. coelicolor* genome reveals homologues of the *min* genes with the exception of *minC*, and a homologue of DivIVA has been analysed (mentioned in section 1.4.1). Another method used for septal placement is known as 'nucleoid occlusion'. This is based on the observation that septal rings only form in areas that contain much less DNA than other areas. Little is known about this process, however recently the discovery of SlmA in *E. coli* (Bernhardt and de Boer 2005) and Noc in *B. subtilis* (Wu and Errington 2004) indicate a mechanism involving binding of proteins to both DNA and FtsZ. The majority of septa in *S. coelicolor* form in the aerial hyphae prior to spore formation, however a few can be found in the vegetative hyphae.

Once the septum has been placed, and the Z – ring formed, a series of proteins join the 'divisome' and complete the process of septal invagination and cell wall synthesis / modification. As with much microbiology, this mechanism has been studied the most in *E. coli*, and so there will be variations between the 'model' scheme and the

situation in *S. coelicolor*. Z-rings are known to be stabilised by two proteins, ZipA and FtsA which anchors FtsZ to the membrane (RayChaudhuri 1999; Pichoff and Lutkenhaus 2005). FtsA also recruits the 'late division' proteins which complete cell division. The late division proteins are all transmembrane proteins containing either a single or multiple transmembrane domains. Of these proteins, FtsEX, FtsN, FtsQ, FtsL and FtsB are of unknown function, however Q, L and B have been shown to assemble into a stable complex (Goehring, Gueiros-Filho, and Beckwith 2005). Little more is known about the other late division proteins – FtsK is involved in chromosome segregation, and it may interact with FtsN to stabilise the divisome (Goehring, Robichon, and Beckwith 2006). AmiC is an amidase required for cell separation in conjunction with penicillin binding proteins (Priyadarshini, Popham, and Young 2006). From the point of view of this thesis, it is the proteins FtsW and FtsI which spark the most interest.

1.7.2 FtsW – like proteins

These proteins are members of the shape, elongation, division and sporulation (SEDS) family, and FtsW contains a series of transmembrane domains (10 in *E. coli*) with varying length loops between the domains. Other proteins that share homology to FtsW are RodA (*E. coli, B. subtilis*), and SpoVE (*B. subtilis*). SpoVE is essential for formation of the *B. subtilis* endospore peptidoglycan, and RodA mutants in *E. coli* and *B. subtilis* assume a spherical shape (Henriques et al. 1998). This evidence indicates a role for RodA as an elongation protein. In lieu of definitive evidence *in vivo* or *in vitro* for the function of FtsW, efforts have concentrated on localisation and mutation studies, many of which have been performed outside of *E. coli* thus allowing comparisons between organisms to be made. It co – localises with FtsZ in *E. coli* (Weiss 2004), *M. tuberculosis* (Datta et al. 2002), *Streptococcus pneumoniae* (Morlot

et al. 2004), and *Mycobacterium smegmatis* (Rajagopalan et al. 2005). This localisation appears to be intimately linked with the formation of septal peptidoglycan, as FtsI (a HMW PBP) and its orthologues in other bacteria also localise to the Z-ring in an FtsW dependent manner (Mercer and Weiss 2002; Morlot et al. 2003). This evidence suggests a role in the stabilisation of the Z-ring, as well as providing a link to the synthesis of septal peptidoglycan. Analysis of the transmembrane domains indicated an interaction between FtsI and the loop between domains 9 and 10 (Pastoret et al. 2004). The genomic arrangement of genes coding FtsW – like proteins also indicates a relationship between themselves and PBPs, as the general genomic arrangement is to find *ftsW* and *ftsI* – like genes clustered with each other (*rodA:pbp2* and *ftsW:ftsI* in *E. coli*). Mutation of *ftsW* in *E. coli* is lethal (Boyle et al. 1997), however in *S. coelicolor* the situation is complicated by the presence of four genes coding for RodA / SpoVE / FtsW - like proteins. One of these genes, SCO3846, is essential (B. Mistry, unpublished results). This multiplicity of FtsW – like proteins in *S. coelicolor* probably reflects the different stages of development, and thus the different kinds of peptidoglycan required for each stage.

1.7.3 The role of penicillin binding proteins

The PBPs have been described in general terms in section 1.6.1, however elucidating the functions of various PBPs is difficult due to the high degree of redundancy between them (Denome et al. 1999). The redundancy poses some serious questions about accepted models of cell wall assembly / modification. It has been shown that a single enzyme, PBP1, is required to manufacture peptidoglycan that resists osmotic stress in *E. coli*. What then of the multi – enzyme complex required for the '3 for 1' model of peptidoglycan assembly? Again, the scaffold model provides an alternative route to the formation of the cell wall that fits with this observation. Of the *E. coli*

PBPs, it is known that PBP2 is specific for cell elongation, while PBP3 is specific for cell division. The other PBPs appear to have a more general role in the synthesis of peptidoglycan. *B. subtilis* has many more PBPs, many of which (PBP2c, PBP2d, SpoVD, PBP3, PBP4b, DacF, PBP4* and PBP5*) are specific for sporulation. *B. subtilis* also has PBPs for general peptidoglycan synthesis as well as some dedicated to septation (Scheffers and Pinho 2005). From the perspective of this thesis, it is *Streptococcus pneumoniae* PBP2x and its orthologues which are most interesting. These proteins belong to the Group B HMW family, and they consist of a C – terminal transmembrane region followed by a PBP dimer region implicated in polymerisation of PBPs (Bertsche et al. 2005). The next domain acts as a transpeptidase, and at the N – terminal lies a pair of 'PASTA' domains (for **P**BP and **s**erine / **t**hreonine kinase **a**ssociated) (Yeats, Finn, and Bateman 2002). Most firmicutes carry one or more orthologues of this protein. The PASTA domains have been shown to weakly bind β – lactam antibiotics (Gordon et al. 2000), and this led to Yeats et. al. speculating on their being located at the site of cell division by sensing a pool of unlinked peptidoglycan. Actinomycetes also possess PBPs with N – terminal PASTA domains, for example PonA2 in *M. tuberculosis,* the transpeptidase of which is believed to catalyse the formation of DAP – DAP linkages between Lipid II subunits (Flores, Parsons, and Pavelka 2005). Homologues to *ponA2* can be found in the *S. coelicolor* genome, and so it is likely that they catalyse similar reactions. The localisation of PBP2x to the cell division septum with FtsZ and FtsW has been shown in *S. pneumoniae* (Morlot et al. 2004), further implicating this PBP in septal peptidoglycan synthesis.

1.8 Signalling in bacteria

1.8.1 C lassical theory of prokaryotic phosphorelays

The division of signalling systems between eukaryotes and prokaryotes was thought to be a clear cut case of the former using phosphorylated serine / threonine / tyrosine residues and the latter using phosphorylated aspartate. This assumption has been overturned, both with the discovery of ser / thr kinases in prokaryotes and the discovery of histidine kinases in eukaryotes (Klumpp and Krieglstein 2002; Klumpp and Krieglstein 2005). The phosporylated histidine system is dependent on a 'two component' arrangement, consisting of a sensor kinase and a response regulator. The kinase autophosphorylates on a histidine residue, then the phosphate group is passed on to a conserved aspartate residue on the response regulator (often a transcription factor) and the required response is elicited (West and Stock 2001). Taking *S. coelicolor* as an example, there are 84 sensor kinases (SK) that interact with 64 response regulators (RR). The majority of genes coding for SKs and RRs are paired together on the chromosome, although there are some 'orphans' which are unpaired. The functions of the SKs and RRs of *S. coelicolor* are mostly unknown; known functions include diverse systems such as osmotic shock, chitinase production and cell wall damage (Hutchings et al. 2004). One of the better characterised is the *vanRS* system coding for a SK – RR pair involved in glycopeptide antibiotic resistance. The sensor kinase (VanS) has a short extracellular domain that either senses glycopeptides or a cell wall intermediate that builds up in response to these antibiotics. The autophosphorylated SK then phosphorylates VanR, the RR, which goes on to induce the rest of the *van* genes coding for glycopeptide resistance (Hong et al. 2004; Hutchings, Hong, and Buttner 2006).

1.8.2 Phospho rylation on serine and threonine residues

The idea of two component systems using autophosphorylating histidine kinases being the exclusive method by which bacteria use phosphosignalling was challenged by the discovery of a kinase that used serine and threonine as the acceptor residues in a manner more reminiscent of eukaryotic systems in *Myxococcus xanthus* (Munoz-Dorado, Inouye, and Inouye 1991). This discovery was rapidly followed by reports of ser / thr kinases and phosphatases in other prokaryotic organisms (Kennelly 2002).

1.8.2.1 Serine and threonine kinases

The ser / thr kinases found in prokaryotes have often been referred to as 'eukaryotic', which is a misnomer arising from the assumptions outlined above. These kinases are in fact of ancient origin, and it is believed that they evolved in the common ancestor well before eukarya, prokarya and archaea pursued their separate evolutionary paths (Leonard, Aravind, and Koonin 1998). The kinases found in bacteria are more diverse than their eukaryotic counterparts, consisting of disparate enzymes that have evolved the ability to phosphorylate ser / thr residues parallel to the canonical 'Hanks' type kinase (Hanks 1988).

The functions of ser / thr kinases are diverse, even in the same organism. The *M. tuberculosis* genome, for example, contains genes coding for 11 ser / thr kinases, PknA – PknL, (Av-Gay and Everett 2000). Table 1 summarises the functional knowledge available about each of these.

Name	Reference(s)	Genomic environment	Function
PknA	(Chaba, Raje, and Chakraborti 2002; Kang et al. 2005)	*pknB* cluster	Regulator of morphological development.
PknB	(Av-Gay, Jamil, and Drews 1999; Yeats, Finn, and Bateman 2002; Grundner, Gay, and Alber 2005; Kang et al. 2005; Dasgupta et al. 2006; Jones and Dyson 2006a)	*pknB* cluster	Regulator of morphological development.
PknD	(Grundner, Gay, and Alber 2005)	*pst* family of phosphate transporters	No knockout phenotype, possibly regulates redundant function. Phosphorylates Rv1747 FHA domain.
PknE	(Grundner, Gay, and Alber 2005)	Proximal to *pknF*	No knockout phenotype, possibly regulates redundant function. Phosphorylates Rv1747 FHA domain.
PknF	(Koul et al. 2001; Molle et al. 2004; Curry et al. 2005; Deol et al. 2005; Grundner, Gay, and Alber 2005)	Proximal to *pknE* and ABC transporter *RV1747*	Regulation of sugar uptake via phosphorylation of ABC transporter.
PknG	(Cowley et al. 2004)	Proximal to glutamate transporter gene *glnH.*	Sensing and regulating glutamine / glutamate levels.
PknH	(Molle et al. 2003; Sharma et al. 2004; Alderwick et al. 2006; Sharma et al. 2006)	Proximal to regulatory gene *embR.*	Virulence regulation, ethambutol resistance via phosphorylation of EmbR.
PknI	(Av-Gay and Everett 2000; Gopalaswamy, Narayanan, and Narayanan 2004)	Proximal to Protein secretion and cell morphology genes *dacB, ffh, ftsY.*	Autophosphorylates, predicted cell division regulation.
PknJ	(Av-Gay and Everett 2000)	Transposon	Function unknown.
PknK	(Av-Gay and Everett 2000)	*luxA* homologue.	Predicted to regulate transcription and secondary metabolism.
PknL	(Av-Gay and Everett 2000)	Transcriptional regulator.	Predicted to regulate transcription.

Table 1: The STPKs of *M. tuberculosis.* Genomic environment and function

There appears to be a correlation between the number of STPKs and complexity of the organism's morphological development, in that microorganisms with complex morphology can have a large number of STPKs, however studied organisms with less involved morphology have not revealed any with a large complement yet. For example, *S. coelicolor* has a predicted 34 STPKs (Petrickova and Petricek 2003), which is dwarfed by the presence of 99 STPKs in *Myxococcus xanthus* as revealed by BLAST searches. This agrees with the idea that increasingly sophisticated morphology requires increasingly sophisticated regulation. Cavalier – Smith took this idea further and posits that eukarya and actinobacteria share a common ancestor (Cavalier-Smith 2002a; Cavalier-Smith 2002b).

Structurally, the STPKs are similar in both prokaryotes and eukaryotes, all possessing the 'Hanks' type structure (Hanks 1988). PknB from *M. tuberculosis* has been characterised structurally, and this is examined in section 4.1. The conserved structure suggests a conserved mode of action for all kinases of this type, with autophosphorylation playing a key role (Ortiz-Lombardia et al. 2003; Young et al. 2003). Autophosphorylation is followed by the transfer of the phosphate to a substrate, and it is the diversity of possible substrates that allows such a wide range of regulatory actions to be performed by STPK molecules. *In vitro* studies have identified 'forkhead associated' (FHA) domains as being the most likely acceptors of the phosphate group (Li et al. 2000; Pallen, Chaudhuri, and Khan 2002; Molle et al. 2003; Grundner, Gay, and Alber 2005; Villarino et al. 2005). It is also worth noting that the bacterial STPKs readily phosphorylate eukaryotic proteins such as myelin basic protein (MBP) *in vitro*, demonstrating a pluralistic approach to their possible

substrates (Av-Gay, Jamil, and Drews 1999). *In vivo* work has been more limited, however currently we know of four substrates, which will be detailed in section 1.9.

1.8.2.2 Serine and threonine phosphatases

Structurally, the ser / thr phosphatases are remarkably similar to each other, and distinct from the tyr phosphatases. They all belong to the PP2C family, and they also share a conserved mode of action, using a metal – bound water molecule as a nucleophile to attack the phosphorus atom in the phosphate group in a S_N2 mechanism (Barford 1996). Curiously, there seems to be little correlation between the number of prokaryotic PP2C phosphatases and the number of STPKs. For example, *Bacillus subtilis* (3 predicted STPKs from BLAST searches) and *M. tuberculosis* (11 STPKs, see above) have a single PP2C. *M. xanthus* (99 STPKs) has four PP2Cs, whilst *S. coelicolor* (34 STPKs (Petrickova and Petricek 2003)) has 95 PP2Cs. It appears that serine / threonine phosphatases in prokaryotes have followed different evolutionary paths than the STPKs, with duplication in the *Streptomyces* being necessary as more complex signalling and sensory pathways evolved (Zhang and Shi 2004). The relative scarcity of phosphatases in *Myxococcus* is still puzzling, as this organism also displays a highly complex morphology requiring a high degree of signalling. It may be that a different method of dephosphorylating p-ser and p-thr is present in this organism.

Genes coding for STPKs are often found proximal to genes coding for ser / thr phosphatases, which would imply a functional relationship (see section 3.2). This functional relationship has been shown for the phosphatase PstP, genetically linked to PknB in *M. tuberculosis* (Boitel et al. 2003; Chopra et al. 2003), and for PrpC linked to PrkC in *B. subtilis* (Gaidenko, Kim, and Price 2002). The above examples were

demonstrated to have an opposite effect to their partner kinases, either dephosphorylating the kinase (PstP – PknB) or the substrate of the kinase (PrkC – PrpC).

1.8.2.3 The role of Forkhead associated (FHA) domains in ser / thr phosphorylation

FHA domains are relatively recent additions to the canon of molecules that are used in a ser / thr signalling cascade. They are found in eukaryotes coupled to ser / thr kinases (Sun et al. 1998; Hammet et al. 2000), as well as other functional domains that are involved in a variety of signalling pathways such as 'forkhead' transcription factors (this is the origin of the domain's name) (Hofmann and Bucher 1995) and mitosis (Elmar Endl 2000). Their function has been predicted to be 'mediators' of ser / thr phosphorylation / dephosphorylation (Li et al. 2000). The discovery of FHA domains in eukaryotes and prokaryotes was simultaneous, which prevented speculation that they were eukaryote specific as happened with STPK molecules (Hofmann and Bucher 1995), and it has been suggested that eukaryotic FHA domains evolved from prokaryotic progenitors (Durocher and Jackson 2002). Their prokaryotic role appears to have an added dimension to that found in eukaryotes. Instead of being purely auxiliary molecules, some proteins are lone agents containing a single FHA domain. It does seem that these lone proteins are in fact analogous to the response regulator found in the histidine kinase system, effecting a response without the aid of a phosphatase molecule. Curr ent evidence for responses effected by bacterial FHA domains is limited, and once again the most information we have comes from the study of *M. tuberculosis*, examples of which are outlined below.

EmbR, the product of gene Rv1267, contains an N - terminal DNA binding domain, a bacterial transcription activation domain and an FHA domain at the C – terminus (Alderwick et al. 2006). This protein has been shown to regulate transcription of the *embCAB* genes, which are integral in resistance to ethambutol, an important antitubercular drug. The STPK PknH has been proven to phosphorylate EmbR, and the phosphorylated species of this protein upregulates the *embCAB* genes and increase resistance to ethambutol (Sharma et al. 2006). Particularly interesting in this case is the upregulation of *pknH* during infection of murine macrophages, coupled with the control of lipoarabinomannan : mannan ratios (important in mycobacterial virulence). This is one example of a STPK / FHA domain pairing that does not appear to need a ser / thr phosphatase for normal function.

Rv1747 is another example of mediation of signal by FHA domains. The protein encoded by this gene is an ABC transporter and containing two FHA domains, and is phosphorylated by PknF, PknB, PknD and PknE *in vitro* (Grundner, Gay, and Alber 2005). It has been suggested that PknF plays a role in glucose transport and cell shape / septation via phosphorylation of the Rv1747 gene product (Molle et al. 2004; Deol et al. 2005). This process appears to be very important for the survival of mycobacteria in mice, as Rv1747 deficient strains failed to grow in the mouse model (Curry et al. 2005).

It is likely that many other examples of STPK – FHA domain interactions will be discovered. Pallen's 2002 survey of available sequences found FHA domains in all major clades of eubacteria, although they have not yet been found in archaebacteria (Pallen, Chaudhuri, and Khan 2002). Currently, there is a notion that FHA domains

are as important, if not more so, than PP2C phosphatases in bacterial ser / thr signalling pathways. Whether this is due to a current fascination with FHA domains to the exclusion of phosphatases, or whether this represents a genuinely surprising shift away from preconceived ideas of signalling pathways remains to be seen.

1.9 PknB – like kinases: The story so far

1.9.1 Pk nB in Actinobacteria

A PknB – like kinase is defined as having a cytoplasmic N – terminus Hanks type STPK domain, followed by a juxtamembrane domain which is linked to a transmembrane domain followed by a number of PASTA domains which lie at the C – terminus and are physically located in the cell wall. The majority of actinobacteria possess just one PknB – like kinase, located in the PknB cluster, although the *Streptomyces* have two (see chapters 3 and 4). PknB from *M. tuberculosis* has been studied intensively due to its essentiality (Sassetti, Boyd, and Rubin 2003; Fernandez et al. 2006), and this has been slightly at the expense of its partner kinase PknA. PknB has been shown to phosphorylate a variety of substrates *in vitro*, and these are listed in table 2.

In vitro substrate	Substrate function	Reference(s)
PknB (autophosphorylation)	Sensing peptidoglycan and signalling.	(Av-Gay, Jamil, and Drews 1999; Yeats, Finn, and Bateman 2002; Boitel et al. 2003; Jones and Dyson 2006a)
MBP	Myelin basic protein. *In vitro* phosphoacceptor for confirmation of kinase activity.	(Av-Gay, Jamil, and Drews 1999)
Rv1747	ABC transporter implicated in virulence.	(Molle et al. 2004; Curry et al. 2005; Grundner, Gay, and Alber 2005)
GarA	Unknown	(Villarino et al. 2005)
Rv0020	Unknown	(Grundner, Gay, and Alber 2005)
Rv1422	Unknown	(Kang et al. 2005)
Wag31	DivIVA orthologue	(Kang et al. 2005)
PBPA	Penicillin binding protein	(Dasgupta et al. 2006)

Table 2: *In vitro* PknB substrates in *M. tuberculosis*. All substrates are phosphorylated on threonine

Of the substrates listed in table 2, Wag31 and Rv1422 have been identified *in vivo* via examination of whole cell extract overexpressing PknB probed with p-ser / thr specific antibodies (Kang et al. 2005). Of the other substrates identified, five contain FHA domains. Villarino et. al. provided a compelling model explaining the importance of specific protein conformation in the interaction of STPKs and FHA domains based on the observation that when presented with tryptic peptide fragments from GarA and MBP that contained the sites identified as PknB phosphorylation targets, no phosphorylation occurred. This has led to the hypothesis that FHA domains interact with the phosphorylated form of STPKs via a pair of autophosphorylated threonines in the activation loop of the enzyme (Villarino et al.

2005). This sits well with site directed mutagenesis experiments that demonstrated phosphorylated thr171 and thr173 are essential for proper function of PknB *in vitro* (Boitel et al. 2003). Another intriguing possibility is that the juxtamembrane region of PknB is involved in the regulation of PknB function. This speculation is based on two phosphorylated threonine residues found in this region, and this pattern of phosphorylation is found in all *M. tuberculosis* STPKs (Duran et al. 2005). Although site directed mutagenesis of these residues did not affect kinase function *in vitro*, there is an argument to suggest a similar mode of regulation to that of PrkC in *B. subtilis* (Madec et al. 2003).

Despite the growing evidence for PknB relaying its phosphosignal via an FHA domain, there is still little evidence for any specific regulatory pathway in which it is involved. The elucidation of this pathway is not helped by the lack of specificity of STPKs when given a choice of substrates to phosphorylate. It can be said that PknB like kinases do play an important role in cell shape and division in mycobacteria, as RNA interference experiments resulted in cells depleted in PknB being longer and thinner. Pleasingly, overexpression of PknB resulted in fatter cells with incomplete septa (Kang et al. 2005). This does suggest that there is a connection between PknB like kinases and regulation of cell shape / division, especially so in light of its genomic location and probable function of the PASTA domains (Jones and Dyson 2006b).

1.9.2 Pk nB in firmicutes

The PknB – like kinases in firmicutes have been studied in a wider range of organisms than is the case in actinobacteria, and this study allows a broader range of functions to be assigned to this class of proteins. The model for the firmicute kinases

is PrkC from *B. subtilis*. The gene encoding this kinase, then called *yloP* was first encountered in 1998 during the sequencing of the genome of this organism. The cognate phosphatase, now known as PrpC coded for by *yloO*, was also correctly identified as being involved in a serine / threonine phosphorylation event (Foulger and Errington 1998). Subsequently, these genes were cloned and characterised. It was found that, like PknB, PrkC autophosphorylated and could phosphorylate MBP. The same investigation also deduced that dimerisation involving the transmembrane and PASTA domains was possible. Mutation of the kinase and the phosphatase resulted in abnormal sporulation and biofilm formation (Madec et al. 2002). The effect of deleting these two genes was further investigated by Gaidenko et. al., and it was found that deletion of *prkC* decreased cell numbers in late stationary phase, whilst deletion of *prpC* increased cell number. The authors also found that a protein phosphorylated on serine present in *prpC* null mutants and absent in *prkC* null mutants corresponded to an elongation factor, EF-G. Following *in vitro* studies showed EF-G being phosphorylated by PrkC and then dephosphorylated by PrpC (Gaidenko, Kim, and Price 2002). The idea of each protein opposing the function of the other has been lent weight by the confirmation of *prkC, prpC* and *yloQ* being transcribed as an operon (Madec et al. 2002; Iwanicki et al. 2005), as cotranscribed genes are often involved in the same process. It is also noteworthy that thus far no FHA domains have been identified in this firmicute system.

Other genes orthologous to *prkC* have also been examined. In *Streptococcus pneumoniae*, the kinase (StkP) has been shown, again, to autophosphorylate and to be dephosphorylated by its cognate phosphatase. An *in vivo* substrate has also been discovered in the form of phosphoglucosamine mutase (GlmM), which plays a crucial

role in the first steps of peptidoglycan synthesis (Novakova et al. 2005). This kinase has also been shown to have an important role in virulence, as virulent strains of *S. pneumoniae* which have had *stkP* deleted show much poorer invasive behaviour than their parental strains (Echenique et al. 2004). Work performed on the *Streptococcus agalactiae* orthologues, Stk1 and Stp1, has confirmed the co – transcription of the kinase and phosphatase, and has identified an inorganic pyrophosphatase as a substrate for the two enzymes (Rajagopal, Clancy, and Rubens 2003). Further work identified the proteins as being involved in the regulation of purine biosynthesis (Rajagopal et al. 2005).

In summary, firmicute PknB – like kinases have similar properties to their actinobacterial counterparts. They autophosphorylate and can be dephosphorylated by their cognate phosphatases. They have the ability to phosphorylate a variety of substrates (although no FHA domain containing substrates have been identified), and this ability has led to different substrates being identified in each species. As in the actinobacteria, the firmicute PknB – like kinases seem to be involved in cell growth and division, however the specific pathway that these proteins regulate is not clear.

1.10 Summary

Streptomyces provide the majority of antibiotics used by mankind, and they are the source of a large number of other useful compounds stemming from their diverse secondary metabolism. The genetic similarity between *S. coelicolor* and *M. tuberculosis* adds a second dimension to the value of studying this organism. Genetically, secondary metabolism and *Streptomyces* development are intimately linked, and they are underpinned by the *whi* and *bld* genes. The ability to sense changes in the external environment is essential for appropriate developmental timing.

Sensing one factor of the external environment can trigger a signalling pathway leading to the correct response to that factor. In the case of PknB – like kinases, it is thought that the PASTA domains sense changes in the state of the bacterial cell wall and then relay a signal to the internal environment, probably via one or more proteins containing FHA domains. What the target of this signal is remains unclear, however it has been shown many times that PknB like kinases are important players in cell morphology and virulence. Circumstantial evidence links PknB – like kinases to cell division, as PBP's can contain PASTA domains and the genes encoding PknB – like kinases are often found in genetic proximity to other cell division genes such as *ftsW*. This link between cell division and sensing the external environment may be a crucial early link in the signalling chain leading to a bacterial cell committing itself to division.

1.11 Introduction References

Alderwick, L. J., V. Molle, L. Kremer, A. J. Cozzone, T. R. Dafforn, G. S. Besra, and K. Futterer. 2006. Molecular structure of EmbR, a response element of Ser/Thr kinase signaling in Mycobacterium tuberculosis 10.1073/pnas.0507766103. PNAS **103**:2558-2563.

Anderson, A., and E. Wellington. 2001. The taxonomy of Streptomyces and related genera Int J Syst Evol Microbiol **51**:797-814.

Anderson, D. E., F. J. Gueiros-Filho, and H. P. Erickson. 2004. Assembly Dynamics of FtsZ Rings in Bacillus subtilis and Escherichia coli and Effects of FtsZ-Regulating Proteins. J. Bacteriol. **186**:5775-5781.

Av-Gay, Y., and M. Everett. 2000. The eukaryotic-like Ser/Thr protein kinases of Mycobacterium tuberculosis. Trends Microbiol **8**:238-244.

Av-Gay, Y., S. Jamil, and S. J. Drews. 1999. Expression and characterization of the Mycobacterium tuberculosis serine/threonine protein kinase PknB. Infect Immun **67**:5676-5682.

Barford, D. 1996. Molecular mechanisms of the protein serine/threonine phosphatases. Trends Biochem Sci **21**:407-412.

Bentley, S. D., K. F. Chater, A. M. Cerdeno-Tarraga, G. L. Challis, N. R. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C. W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C. H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Neil, E. Rabbinowitsch, M. A. Rajandream, K.

Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B. G. Barrell, J. Parkhill, and D. A. Hopwood. 2002. Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). Nature **417**:141-147.

Bentley, S. D., M. Maiwald, L. D. Murphy, M. J. Pallen, C. A. Yeats, L. G. Dover, H. T. Norbertczak, G. S. Besra, M. A. Quail, and D. E. Harris. 2003. Sequencing and analysis of the genome of the Whipple's disease bacterium Tropheryma whipplei. The Lancet **361**:637-644.

Bernhardt, T. G., and P. A. J. de Boer. 2005. SlmA, a Nucleoid-Associated, FtsZ Binding Protein Required for Blocking Septal Ring Assembly over Chromosomes in E. coli. Molecular Cell **18**:555-564.

Bertsche, U., E. Breukink, T. Kast, and W. Vollmer. 2005. In Vitro Murein (Peptidoglycan) Synthesis by Dimers of the Bifunctional Transglycosylase-Transpeptidase PBP1B from Escherichia coli
10.1074/jbc.M508646200. J. Biol. Chem. **280**:38096-38101.

Boitel, B., M. Ortiz-Lombardia, R. Duran, F. Pompeo, S. T. Cole, C. Cervenansky, and P. M. Alzari. 2003. PknB kinase activity is regulated by phosphorylation in two Thr residues and dephosphorylation by PstP, the cognate phospho-Ser/Thr phosphatase, in Mycobacterium tuberculosis. Mol Microbiol **49**:1493-1508.

Boyle, D. S., M. M. Khattar, S. G. Addinall, J. Lutkenhaus, and W. D. Donachie. 1997. ftsW is an essential cell-division gene in Escherichia coli. Mol Microbiol **24**:1263-1273.

Cavalier-Smith, T. 2002a. The neomuran origin of archaebacteria, the negibacterial root of the universal tree and bacterial megaclassification. Int J Syst Evol Microbiol **52**:7-76.

Cavalier-Smith, T. 2002b. The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. Int J Syst Evol Microbiol **52**:297-354.

Chaba, R., M. Raje, and P. K. Chakraborti. 2002. Evidence that a eukaryotic-type serine/threonine protein kinase from Mycobacterium tuberculosis regulates morphological changes associated with cell division. Eur J Biochem **269**:1078-1085.

Chater, K. F. 2001. Regulation of sporulation in Streptomyces coelicolor A3(2): a checkpoint multiplex? Curr Opin Microbiol **4**:667-673.

Chater, K. F., and S. Horinouchi. 2003. Signalling early developmental events in two highly diverged Streptomyces species
doi:10.1046/j.1365-2958.2003.03476.x. Molecular Microbiology **48**:9-15.

Chen, C. W., C. H. Huang, H. H. Lee, H. H. Tsai, and R. Kirby. 2002. Once the circle has been broken: dynamics and evolution of Streptomyces chromosomes. Trends Genet **18**:522-529.

Chen, Y., and H. P. Erickson. 2005. Rapid in Vitro Assembly Dynamics and Subunit Turnover of FtsZ Demonstrated by Fluorescence Resonance Energy Transfer
10.1074/jbc.M500895200. J. Biol. Chem. **280**:22549-22554.

Chopra, P., B. Singh, R. Singh, R. Vohra, A. Koul, L. S. Meena, H. Koduri, M. Ghildiyal, P. Deol, T. K. Das, A. K. Tyagi, and Y. Singh. 2003. Phosphoprotein phosphatase of Mycobacterium tuberculosis dephosphorylates serine-threonine kinases PknA and PknB. Biochem Biophys Res Commun **311**:112-120.

Claessen, D., W. de Jong, L. Dijkhuizen, and H. A. B. Wosten. 2006. Regulation of Streptomyces development: reach for the sky! Trends in Microbiology **14**:313-319.

Claessen, D., I. Stokroos, H. J. Deelstra, N. A. Penninga, C. Bormann, J. A. Salas, L. Dijkhuizen, and H. A. B. Wosten. 2004. The formation of the rodlet layer of streptomycetes is the result of the interplay between rodlins and chaplins doi:10.1111/j.1365-2958.2004.04143.x. Molecular Microbiology **53**:433-443.

Cowley, S., M. Ko, N. Pick, R. Chow, K. J. Downing, B. G. Gordhan, J. C. Betts, V. Mizrahi, D. A. Smith, R. W. Stokes, and Y. Av-Gay. 2004. The Mycobacterium tuberculosis protein serine/threonine kinase PknG is linked to cellular glutamate/glutamine levels and is important for growth in vivo. Molecular Microbiology **52**:1691-1702.

Curry, J. M., R. Whalan, D. M. Hunt, K. Gohil, M. Strom, L. Rickman, M. J. Colston, S. J. Smerdon, and R. S. Buxton. 2005. An ABC Transporter Containing a Forkhead-Associated Domain Interacts with a Serine-Threonine Protein Kinase and Is Required for Growth of Mycobacterium tuberculosis in Mice. Infect. Immun. **73**:4471-4477.

Dai, K., and J. Lutkenhaus. 1991. ftsZ is an essential cell division gene in Escherichia coli. J Bacteriol **173**:3500-3506.

Dasgupta, A., P. Datta, M. Kundu, and J. Basu. 2006. The serine/threonine kinase PknB of Mycobacterium tuberculosis phosphorylates PBPA, a penicillin-binding protein required for cell division 10.1099/mic.0.28630-0. Microbiology **152**:493-504.

Datta, P., A. Dasgupta, S. Bhakta, and J. Basu. 2002. Interaction between FtsZ and FtsW of Mycobacterium tuberculosis. J Biol Chem **277**:24983-24987.

Del Sol, R., J. G. Mullins, N. Grantcharova, K. Flardh, and P. Dyson. 2006. Influence of CrgA on Assembly of the Cell Division Protein FtsZ during Development of Streptomyces coelicolor. J Bacteriol **188**:1540-1550.

Del Sol, R., A. Pitman, P. Herron, and P. Dyson. 2003. The product of a developmental gene, crgA, that coordinates reproductive growth in Streptomyces belongs to a novel family of small actinomycete-specific proteins. J Bacteriol **185**:6678-6685.

Denome, S. A., P. K. Elf, T. A. Henderson, D. E. Nelson, and K. D. Young. 1999. Escherichia coli Mutants Lacking All Possible Combinations of Eight Penicillin Binding Proteins: Viability, Characteristics, and Implications for Peptidoglycan Synthesis. J. Bacteriol. **181**:3981-3993.

Deol, P., R. Vohra, A. K. Saini, A. Singh, H. Chandra, P. Chopra, T. K. Das, A. K. Tyagi, and Y. Singh. 2005. Role of Mycobacterium tuberculosis Ser/Thr Kinase PknF: Implications in Glucose Transport and Cell Division. J. Bacteriol. **187**:3415-3420.

Deutscher, J., and M. H. Saier, Jr. 2005. Ser/Thr/Tyr protein phosphorylation in bacteria - for long time neglected, now well established. J Mol Microbiol Biotechnol **9**:125-131.

Dmitriev, B. A., F. V. Toukach, K. J. Schaper, O. Holst, E. T. Rietschel, and S. Ehlers. 2003. Tertiary structure of bacterial murein: the scaffold model. J Bacteriol **185**:3458-3468.

Duran, R., A. Villarino, M. Bellinzoni, A. Wehenkel, P. Fernandez, B. Boitel, S. T. Cole, P. M. Alzari, and C. Cervenansky. 2005. Conserved autophosphorylation pattern in activation loops and juxtamembrane regions of

Mycobacterium tuberculosis Ser/Thr protein kinases. Biochemical and Biophysical Research Communications **333**:858-867.
Durocher, D., and S. P. Jackson. 2002. The FHA domain. FEBS Lett **513**:58-66.
Echenique, J., A. Kadioglu, S. Romao, P. W. Andrew, and M.-C. Trombe. 2004. Protein Serine/Threonine Kinase StkP Positively Controls Virulence and Competence in Streptococcus pneumoniae. Infect. Immun. **72**:2434-2437.
Elliot, M. A., M. J. Bibb, M. J. Buttner, and B. K. Leskiw. 2001. BldD is a direct regulator of key developmental genes in Streptomyces coelicolor A3(2) doi:10.1046/j.1365-2958.2001.02387.x. Molecular Microbiology **40**:257-269.
Elmar Endl, J. G. 2000. Posttranslational modifications of the KI-67 protein coincide with two major checkpoints during mitosis. Journal of Cellular Physiology **182**:371-380.
Fernandez, P., B. Saint-Joanis, N. Barilone, M. Jackson, B. Gicquel, S. T. Cole, and P. M. Alzari. 2006. The Ser/Thr Protein Kinase PknB Is Essential for Sustaining Mycobacterial Growth. J. Bacteriol. **188**:7778-7784.
Flardh, K. 2003. Essential role of DivIVA in polar growth and morphogenesis in Streptomyces coelicolor A3(2). Molecular Microbiology **49**:1523-1536.
Flardh, K., K. C. Findlay, and K. F. Chater. 1999. Association of early sporulation genes with suggested developmental decision points in Streptomyces coelicolor A3(2). Microbiology **145 (Pt 9)**:2229-2243.
Flores, A. R., L. M. Parsons, and M. S. Pavelka, Jr. 2005. Characterization of Novel Mycobacterium tuberculosis and Mycobacterium smegmatis Mutants Hypersusceptible to {beta}-Lactam Antibiotics. J. Bacteriol. **187**:1892-1900.
Foulger, D., and J. Errington. 1998. A 28 kbp segment from the spoVM region of the Bacillus subtilis 168 genome. Microbiology **144 (Pt 3)**:801-805.
Fu, X., Y.-L. Shih, Y. Zhang, and L. I. Rothfield. 2001. The MinE ring required for proper placement of the division site is a mobile structure that changes its cellular location during the Escherichia coli division cycle 10.1073/pnas.031549298. PNAS **98**:980-985.
Gaidenko, T. A., T. J. Kim, and C. W. Price. 2002. The PrpC serine-threonine phosphatase and PrkC kinase have opposing physiological roles in stationary-phase Bacillus subtilis cells. J Bacteriol **184**:6109-6114.
Goehring, N. W., and J. Beckwith. 2005. Diverse Paths to Midcell: Assembly of the Bacterial Cell Division Machinery. Current Biology **15**:R514-R526.
Goehring, N. W., F. Gueiros-Filho, and J. Beckwith. 2005. Premature targeting of a cell division protein to midcell allows dissection of divisome assembly in Escherichia coli 10.1101/gad.1253805. Genes Dev. **19**:127-137.
Goehring, N. W., C. Robichon, and J. Beckwith. 2006. A role for the non-essential N-terminus of FtsN in divisome assembly. J. Bacteriol.:JB.00992-00906.
Gopalaswamy, R., P. R. Narayanan, and S. Narayanan. 2004. Cloning, overexpression, and characterization of a serine/threonine protein kinase pknI from Mycobacterium tuberculosis H37Rv. Protein Expression and Purification **36**:82-89.
Gordon, E., N. Mouz, E. Duee, and O. Dideberg. 2000. The crystal structure of the penicillin-binding protein 2x from Streptococcus pneumoniae and its acyl-enzyme form: implication in drug resistance. Journal of Molecular Biology **299**:477-485.

Grundner, C., L. M. Gay, and T. Alber. 2005. Mycobacterium tuberculosis serine/threonine kinases PknB, PknD, PknE, and PknF phosphorylate multiple FHA domains. Protein Sci **14**:1918-1921.

Hammet, A., B. L. Pike, K. I. Mitchelhill, T. Teh, B. Kobe, C. M. House, B. E. Kemp, and J. Heierhorst. 2000. FHA domain boundaries of the dun1p and rad53p cell cycle checkpoint kinases. FEBS Lett **471**:141-146.

Hanks, S., S K|Quinn,AM,A M|Hunter,T,T. 1988. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. Science **241**:42-52.

Henriques, A. O., P. Glaser, P. J. Piggot, and C. P. Moran Jr. 1998. Control of cell shape and elongation by the rodA gene in Bacillus subtilis. Molecular Microbiology **28**:235-247.

Hesketh, A., J. Sun, and M. Bibb. 2001. Induction of ppGpp synthesis in Streptomyces coelicolor A3(2) grown under conditions of nutritional sufficiency elicits actII-ORF4 transcription and actinorhodin biosynthesis. Molecular Microbiology **39**:136-144.

Hofmann, K., and P. Bucher. 1995. The FHA domain: a putative nuclear signalling domain found in protein kinases and transcription factors. Trends Biochem Sci **20**:347-349.

Holtje, J. V. 1998. Growth of the stress-bearing and shape-maintaining murein sacculus of Escherichia coli. Microbiol Mol Biol Rev **62**:181-203.

Hong, H.-J., M. I. Hutchings, J. M. Neu, G. D. Wright, M. S. B. Paget, and M. J. Buttner. 2004. Characterization of an inducible vancomycin resistance system in Streptomyces coelicolor reveals a novel gene (vanK) required for drug resistance
doi:10.1111/j.1365-2958.2004.04032.x. Molecular Microbiology **52**:1107-1121.

Hutchings, M. I., H.-J. Hong, and M. J. Buttner. 2006. The vancomycin resistance VanRS two-component signal transduction system of Streptomyces coelicolor
doi:10.1111/j.1365-2958.2005.04953.x. Molecular Microbiology **59**:923-935.

Hutchings, M. I., P. A. Hoskisson, G. Chandra, and M. J. Buttner. 2004. Sensing and responding to diverse extracellular signals? Analysis of the sensor kinases and response regulators of Streptomyces coelicolor A3(2)
10.1099/mic.0.27181-0. Microbiology **150**:2795-2806.

Ikeda, H., J. Ishikawa, A. Hanamoto, M. Shinose, H. Kikuchi, T. Shiba, Y. Sakaki, M. Hattori, and S. Omura. 2003. Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis. Nat Biotechnol **21**:526-531.

Iwanicki, A., K. Hinc, S. Seror, G. WÄ™grzyn, and M. Obuchowski. 2005. Transcription in the prpC-yloQ region in Bacillus subtilis. Archives of Microbiology **183**:421-430.

Jones, G., and P. Dyson. 2006a. Evolution of transmembrane protein kinases implicated in coordinating remodelling of Gram-positive peptidoglycan: inside versus outside. J. Bacteriol.:JB.00800-00806.

Jones, G., and P. Dyson. 2006b. Evolution of Transmembrane Protein Kinases Implicated in Coordinating Remodeling of Gram-Positive Peptidoglycan: Inside versus Outside. J. Bacteriol. **188**:7470-7476.

Kang, C.-M., D. W. Abbott, S. T. Park, C. C. Dascher, L. C. Cantley, and R. N. Husson. 2005. The Mycobacterium tuberculosis serine/threonine kinases PknA and PknB: substrate identification and regulation of cell shape. Genes Dev.:gad.1311105.

Kelemen, G. H., P. Brian, K. Flardh, L. Chamberlin, K. F. Chater, and M. J. Buttner. 1998. Developmental Regulation of Transcription of whiE, a Locus Specifying the Polyketide Spore Pigment in Streptomyces coelicolor A3(2). J. Bacteriol. **180**:2515-2521.

Kennelly, P. J. 2002. Protein kinases and protein phosphatases in prokaryotes: a genomic perspective. FEMS Microbiol Lett **206**:1-8.

Klumpp, S., and J. Krieglstein. 2005. Reversible phosphorylation of histidine residues in vertebrate proteins. Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics **1754**:291-295.

Klumpp, S., and J. Krieglstein. 2002. Phosphorylation and dephosphorylation of histidine residues in proteins

Eur J Biochem **269**:1067-1071.

Koul, A., A. Choidas, A. K. Tyagi, K. Drlica, Y. Singh, and A. Ullrich. 2001. Serine/threonine protein kinases PknF and PknG of Mycobacterium tuberculosis: characterization and localization. Microbiology **147**:2307-2314.

Leonard, C. J., L. Aravind, and E. V. Koonin. 1998. Novel families of putative protein kinases in bacteria and archaea: evolution of the "eukaryotic" protein kinase superfamily. Genome Res **8**:1038-1047.

Li, J., G. Lee, S. Van Doren, and J. Walker. 2000. The FHA domain mediates phosphoprotein interactions. J Cell Sci **113**:4143-4149.

Lutkenhaus, J. F., H. Wolf-Watz, and W. D. Donachie. 1980. Organization of genes in the ftsA-envA region of the Escherichia coli genetic map and identification of a new fts locus (ftsZ). J Bacteriol **142**:615-620.

Madec, E., A. Laszkiewicz, A. Iwanicki, M. Obuchowski, and S. Seror. 2002. Characterization of a membrane-linked Ser/Thr protein kinase in Bacillus subtilis, implicated in developmental processes. Mol Microbiol **46**:571-586.

Madec, E., A. Stensballe, S. Kjellstrom, L. Cladiere, M. Obuchowski, O. N. Jensen, and S. J. Seror. 2003. Mass Spectrometry and Site-directed Mutagenesis Identify Several Autophosphorylated Residues Required for the Activity of PrkC, a Ser/Thr Kinase from Bacillus subtilis. Journal of Molecular Biology **330**:459-472.

Magnusson, L. U., A. Farewell, and T. Nystrom. 2005. ppGpp: a global regulator in Escherichia coli. Trends in Microbiology **13**:236-242.

Massidda, O., D. Anderluzzi, L. Friedli, and G. Feger. 1998. Unconventional organization of the division and cell wall gene cluster of Streptococcus pneumoniae. Microbiology **144**:3069-3078.

McCormick, J. R., E. P. Su, A. Driks, and R. Losick. 1994. Growth and viability of Streptomyces coelicolor mutant for the cell division gene ftsZ. Mol Microbiol **14**:243-254.

Mercer, K. L. N., and D. S. Weiss. 2002. The Escherichia coli Cell Division Protein FtsW Is Required To Recruit Its Cognate Transpeptidase, FtsI (PBP3), to the Division Site. J. Bacteriol. **184**:904-912.

Meroueh, S. O., K. Z. Bencze, D. Hesek, M. Lee, J. F. Fisher, T. L. Stemmler, and S. Mobashery. 2006. Three-dimensional structure of the bacterial cell wall peptidoglycan

10.1073/pnas.0510182103. PNAS **103**:4404-4409.

Mingorance, J., J. Tamames, and M. Vicente. 2004. Genomic channeling in bacterial cell division. J Mol Recognit **17**:481-487.

Molle, V., L. Kremer, C. Girard-Blanc, G. S. Besra, A. J. Cozzone, and J. F. Prost. 2003. An FHA phosphoprotein recognition domain mediates protein EmbR

phosphorylation by PknH, a Ser/Thr protein kinase from Mycobacterium tuberculosis. Biochemistry **42**:15300-15309.

Molle, V., D. Soulat, J. M. Jault, C. Grangeasse, A. J. Cozzone, and J. F. Prost. 2004. Two FHA domains on an ABC transporter, Rv1747, mediate its phosphorylation by PknF, a Ser/Thr protein kinase from Mycobacterium tuberculosis. FEMS Microbiol Lett **234**:215-223.

Morlot, C., M. Noirclerc-Savoye, A. Zapun, O. Dideberg, and T. Vernet. 2004. The d,d-carboxypeptidase PBP3 organizes the division process of Streptococcus pneumoniae. Molecular Microbiology **51**:1641-1648.

Morlot, C., A. Zapun, O. Dideberg, and T. Vernet. 2003. Growth and division of Streptococcus pneumoniae: localization of the high molecular weight penicillin-binding proteins during the cell cycle. Molecular Microbiology **50**:845-855.

Munoz-Dorado, J., S. Inouye, and M. Inouye. 1991. A gene encoding a protein serine/threonine kinase is required for normal development of M. xanthus, a gram-negative bacterium. Cell **67**:995-1006.

Nguyen, K. T., J. M. Willey, L. D. Nguyen, L. T. Nguyen, P. H. Viollier, and C. J. Thompson. 2002. A central regulator of morphological differentiation in the multicellular bacterium Streptomyces coelicolor. Mol Microbiol **46**:1223-1238.

Novakova, L., L. Saskova, P. Pallova, J. Janecek, J. Novotna, A. Ulrych, J. Echenique, M.-C. Trombe, and P. Branny. 2005. Characterization of a eukaryotic type serine/threonine protein kinase and protein phosphatase of Streptococcus pneumoniae and identification of kinase substrates. FEBS Journal **272**:1243-1254.

Ohnishi, Y., H. Yamazaki, J. Y. Kato, A. Tomono, and S. Horinouchi. 2005. AdpA, a central transcriptional regulator in the A-factor regulatory cascade that leads to morphological development and secondary metabolism in Streptomyces griseus. Biosci Biotechnol Biochem **69**:431-439.

Ortiz-Lombardia, M., F. Pompeo, B. Boitel, and P. M. Alzari. 2003. Crystal structure of the catalytic domain of the PknB serine/threonine kinase from Mycobacterium tuberculosis. J Biol Chem **278**:13094-13100.

Pallen, M., R. Chaudhuri, and A. Khan. 2002. Bacterial FHA domains: neglected players in the phospho-threonine signalling game? Trends Microbiol **10**:556-563.

Pastoret, S., C. Fraipont, T. den Blaauwen, B. Wolf, M. E. G. Aarsman, A. Piette, A. Thomas, R. Brasseur, and M. Nguyen-Disteche. 2004. Functional Analysis of the Cell Division Protein FtsW of Escherichia coli. J. Bacteriol. **186**:8370-8379.

Petrickova, K., and M. Petricek. 2003. Eukaryotic-type protein kinases in Streptomyces coelicolor: variations on a common theme. Microbiology **149**:1609-1621.

Pichoff, S., and J. Lutkenhaus. 2005. Tethering the Z ring to the membrane through a conserved membrane targeting sequence in FtsA
doi:10.1111/j.1365-2958.2005.04522.x. Molecular Microbiology **55**:1722-1734.

Priyadarshini, R., D. L. Popham, and K. D. Young. 2006. Daughter Cell Separation by Penicillin-Binding Proteins and Peptidoglycan Amidases in Escherichia coli. J. Bacteriol. **188**:5345-5355.

Rajagopal, L., A. Clancy, and C. E. Rubens. 2003. A Eukaryotic Type Serine/Threonine Kinase and Phosphatase in Streptococcus agalactiae

Reversibly Phosphorylate an Inorganic Pyrophosphatase and Affect Growth, Cell Segregation, and Virulence
10.1074/jbc.M212747200. J. Biol. Chem. **278**:14429-14441.
Rajagopal, L., A. Vo, A. Silvestroni, and C. E. Rubens. 2005. Regulation of purine biosynthesis by a eukaryotic-type kinase in Streptococcus agalactiae. Molecular Microbiology **56**:1329-1346.
Rajagopalan, M., E. Maloney, J. Dziadek, M. Poplawska, H. Lofton, A. Chauhan, and M. V. Madiraju. 2005. Genetic evidence that mycobacterial FtsZ and FtsW proteins interact, and colocalize to the division site in Mycobacterium smegmatis. FEMS Microbiol Lett **250**:9-17.
Raoult, D., H. Ogata, S. Audic, C. Robert, K. Suhre, M. Drancourt, and J. M. Claverie. 2003. Tropheryma whipplei Twist: a human pathogenic Actinobacteria with a reduced genome. Genome Res **13**:1800-1809.
RayChaudhuri, D. 1999. ZipA is a MAP-Tau homolog and is essential for structural integrity of the cytokinetic FtsZ ring during bacterial cell division. Embo J **18**:2372-2383.
Ryding, N. J., M. J. Bibb, V. Molle, K. C. Findlay, K. F. Chater, and M. J. Buttner. 1999. New Sporulation Loci in Streptomyces coelicolor A3(2). J. Bacteriol. **181**:5419-5425.
Ryding, N. J., G. H. Kelemen, C. A. Whatling, K. Flardh, M. J. Buttner, and K. F. Chater. 1998. A developmentally regulated gene encoding a repressor-like protein is essential for sporulation in Streptomyces coelicolor A3(2)
doi:10.1046/j.1365-2958.1998.00939.x. Molecular Microbiology **29**:343-357.
Saito, N., J. Xu, T. Hosaka, S. Okamoto, H. Aoki, M. J. Bibb, and K. Ochi. 2006. EshA Accentuates ppGpp Accumulation and Is Conditionally Required for Antibiotic Production in Streptomyces coelicolor A3(2). J. Bacteriol. **188**:4952-4961.
Sassetti, C. M., D. H. Boyd, and E. J. Rubin. 2003. Genes required for mycobacterial growth defined by high density mutagenesis. Mol Microbiol **48**:77-84.
Scheffers, D.-J., L. J. F. Jones, and J. Errington. 2004. Several distinct localization patterns for penicillin-binding proteins in Bacillus subtilis
doi:10.1046/j.1365-2958.2003.03854.x. Molecular Microbiology **51**:749-764.
Scheffers, D.-J., and M. G. Pinho. 2005. Bacterial Cell Wall Synthesis: New Insights from Localization Studies
10.1128/MMBR.69.4.585-607.2005. Microbiol. Mol. Biol. Rev. **69**:585-607.
Sharma, K., H. Chandra, P. K. Gupta, M. Pathak, A. Narayan, L. S. Meena, R. C. D'Souza, P. Chopra, S. Ramachandran, and Y. Singh. 2004. PknH, a transmembrane Hank's type serine/threonine kinase from Mycobacterium tuberculosis is differentially expressed under stress conditions. FEMS Microbiol Lett **233**:107-113.
Sharma, K., M. Gupta, M. Pathak, N. Gupta, A. Koul, S. Sarangi, R. Baweja, and Y. Singh. 2006. Transcriptional Control of the Mycobacterial embCAB Operon by PknH through a Regulatory Protein, EmbR, In Vivo. J. Bacteriol. **188**:2936-2944.
Spratt, B. G. 1975. Distinct penicillin binding proteins involved in the division, elongation, and shape of Escherichia coli K12. Proc Natl Acad Sci U S A **72**:2999-3003.
Sun, J., A. Hesketh, and M. Bibb. 2001. Functional Analysis of relA and rshA, Two relA/spoT Homologues of Streptomyces coelicolor A3(2). J. Bacteriol. **183**:3488-3498.

Sun, Z., J. Hsiao, D. S. Fay, and D. F. Stern. 1998. Rad53 FHA domain associated with phosphorylated Rad9 in the DNA damage checkpoint. Science **281**:272-274.

T. Kieser, M. J. B., M.J. Buttner, K.F. Chater, D.A. Hopwood. 2000. Practical Streptomyces Genetics. John Innes Foundation.

Takano, E. 2006. [gamma]-Butyrolactones: Streptomyces signalling molecules regulating antibiotic production and differentiation. Current Opinion in Microbiology **9**:287-294.

van Heijenoort, J. 2001. Formation of the glycan chains in the synthesis of bacterial peptidoglycan. Glycobiology **11**:25R-36R.

Vicente, M., K. F. Chater, and V. de Lorenzo. 1999. Bacterial transcription factors involved in global regulation. Molecular Microbiology **33**:8-17.

Villarino, A., R. Duran, A. Wehenkel, P. Fernandez, P. England, P. Brodin, S. T. Cole, U. Zimny-Arndt, P. R. Jungblut, C. Cervenansky, and P. M. Alzari. 2005. Proteomic Identification of M. tuberculosis Protein Kinase Substrates: PknB Recruits GarA, a FHA Domain-containing Protein, Through Activation Loop-mediated Interactions. Journal of Molecular Biology **350**:953-963.

Volff, J.-N., and J. Altenbuchner. 1998. Genetic instability of the Streptomyces chromosome. Molecular Microbiology **27**:239-246.

Vollmer, W., and J.-V. Holtje. 2004. The Architecture of the Murein (Peptidoglycan) in Gram-Negative Bacteria: Vertical Scaffold or Horizontal Layer(s)? J. Bacteriol. **186**:5978-5987.

Watve, M. G., R. Tickoo, M. M. Jog, and B. D. Bhole. 2001. How many antibiotics are produced by the genus Streptomyces? Arch Microbiol **176**:386-390.

Weaver, D., N. Karoonuthaisiri, H.-H. Tsai, C.-H. Huang, M.-L. Ho, S. Gai, K. G. Patel, J. Huang, S. N. Cohen, D. A. Hopwood, C. W. Chen, and C. M. Kao. 2004. Genome plasticity in Streptomyces: identification of 1 Mb TIRs in the S. coelicolor A3(2) chromosome
doi:10.1111/j.1365-2958.2003.03920.x. Molecular Microbiology **51**:1535-1550.

Weiss, D. S. 2004. Bacterial cell division and the septal ring
doi:10.1111/j.1365-2958.2004.04283.x. Molecular Microbiology **54**:588-597.

West, A. H., and A. M. Stock. 2001. Histidine kinases and response regulator proteins in two-component signaling systems. Trends in Biochemical Sciences **26**:369-376.

Wu, L. J., and J. Errington. 2004. Coordination of Cell Division and Chromosome Segregation by a Nucleoid Occlusion Protein in Bacillus subtilis. Cell **117**:915-925.

Yeats, C., R. D. Finn, and A. Bateman. 2002. The PASTA domain: a beta-lactam-binding domain. Trends Biochem Sci **27**:438.

Young, K. D. 2001. Approaching the physiological functions of penicillin-binding proteins in Escherichia coli. Biochimie **83**:99-102.

Young, K. D. 2006. Too many strictures on structure. Trends in Microbiology **14**:155-156.

Young, T. A., B. Delagoutte, J. A. Endrizzi, A. M. Falick, and T. Alber. 2003. Structure of Mycobacterium tuberculosis PknB supports a universal activation mechanism for Ser/Thr protein kinases. Nat Struct Biol **10**:168-174.

Zhang, W., and L. Shi. 2004. Evolution of the PPM-family protein phosphatases in Streptomyces: duplication of catalytic domain and lateral recruitment of additional sensory domains
10.1099/mic.0.27480-0. Microbiology **150**:4189-4197.

2 Materials and methods

2.1 *Strains used*

Strain	Genotype	Source
E. coli JM109	F' *traD36 proA+B+ lacIq* Δ*(lacZ)M15/* Δ*(lac-proAB) glnV44 e14-gyrA96 recA1 relA1endA1 thi hsdR17*	Promega Corp.
E. coli BL21 DE3	*lon-, ompT-*	Novagen
E. coli Novablue	Tet^R	Novagen
E. coli ET12567	*Dam13*::Tn*9 dcm6 hsdM hsdR recF143* 16 zjj*201* ::Tn*10 galK2 galT22 ara14 lacY1 xyl5 leuB6 thi1 tonA31 rpsL136 hisG4 tsx78 mtli glnV44*, containing the nontransmissible *oriT* mobilizing plasmid, pUZ8002	(Flett, Mersinias, and Smith 1997)
S. coelicolor A3(2) M145	Prototrophic $SCP1^-$ $SCP2^-$ Pgl^+	NCIMB
S. avermitilis 12804	Wild Type	NCIMB
DSCO3848	M145::tn5052	This work
DSCO3845	M145::tn5062::pSHP-	This work
DSCO3843/44	M145::tn5066	This work

DSAV4338	12804::tn5066	This work

2.2 *Plasmids / Cosmids / transposons used*

Plasmid	Features	Use	Source
pQM5066	Contains tn5066, Amp^R, Hyg^R	Source of tn5066	R. Del Sol
pQM5062	Contains tn5062, Amp^R, Apr^R	Source of tn5062	R. Del Sol
pRKin1	High copy number, pUWL219 derivative. Apr^R, Amp^R, Tsr^R.	Overexpression of SCO3848	R. Del Sol
SCH69.2.F01	Supercos derivative containing tn5062 disruption of SCO3848. Apr^R, Amp^R.	Construction of DSCO3848.	A. Bishop
pLK1	pIJ2925 derivative, contains SCO3848. Amp^R.	Site directed mutagenesis.	R. Del Sol
pLK2	pAlter derivative, contains SCO3848. Amp^R.	Sub cloning of SCO3848	R. Del Sol
pSH152	pSET152 derivative, Hyg^R.	Integrative plasmid used for complementation of Apr^R disruptants.	R. Del Sol
pSC3848	pSH152 derivative containing SCO3848, Hyg^R.	Complementation of DSCO3848.	R. Del Sol
pME6	Invitrogen pCR2.1-TOPO	Subcloning and transposon	A. Bishop

	derivative, Km^R, Amp^R, *lacZ*, pUC ori, F1 ori.	mutagenesis.	
pHit-K	pLK1 after site directed mutagenesis to introduce *Nde*I site at start of SCO3848.	Source of kinase gene for overexpression in *E. coli.*	This work
pHit-K2	pHit-K after site directed mutagenesis to introduce *Nde*I site in transmembrane region of SCO3848	Source of kinase domain band for overexpression in *E. coli.*	This work
pMEK2	pME6 derivative containing *S. avermitilis* kinase fragment.	Transposon mutagenesis of SAV4338.	This work.
pET16k	Km^R version of pET16b. Contains N – terminal His tag, pBR322 origin, T7 promoter and terminator, *lacI* sequence.	IPTG inducible overexpression vector.	Novagen
pETK1	pET16k containing SCO3848 band from pHit-K.	Overexpression of His tagged SCO3848.	This work
pETK2	PET16k derivative containing SCO3848 kinase domain from pHit-K2.	Overexpression of kinase domain only.	This work
pMEF	PME6 derivative containing	Construction of DSCO3845 and	This work

	7kb SCO3843 – SCO3847 fragment from B. Mistry's pRCLU2	complementation of DSCO3843/44.	
pMEFP-	pMEF digested *Aat*II and re – ligated for in – frame deletion of SCO3845.	Construction of pSHP-.	This work
pSHP-	pSH152 derivative containing SCO3845- fragment from pMEFP-.	Construction of DSCO3845.	This work
SCh69.1.C06	Supercos derivative carrying tn5062 disruption of SCO3845.	Construction of DSCO3845.	L. Fernandez
pMER	pME6 derivative containing SCO3844 deletion.	Construction of pMERL	This work
pMERL	pMER derivative containing SCO3843 and SCO3844 deletions.	Construction of pMERL66	This work
pMERL66	pMERL derivative containing tn5066 between SCO3843 and SCO3844.	Construction of DSCO3843/44	This work
pIJFL	pIJ2925 derivative containing part SCO3841 – part SCO3843	Construction of pMERL.	This work

	fragment.		
pSEF2	pSET152 derivative containing 2.68kb *Pvu*II / *Eco*RV fragment from pMEF (SCO3843 + SCO3844).	Complementation of DSCO3843/44	This work
pLPK1	PIJ5971 derivative containing kinase promoter region fused to *luxA/B* system.	*lux* reporter analysis of SCO3848 expression.	R. Del Sol

Table 2.2: Plasmids and cosmids used in this study.

2.3 *Growth media*

Name	**Composition**	**Amount (g/l dH_2O unless otherwise stated)**
LB (Luria Bertani) Medium	Tryptone	10
	Yeast Extract	5
	NaCl	5
	Glucose	1 (optional)
	adjust pH to 7.0 with NaOH	
	Agar	10
2XYT	Tryptone	16
	Yeast Extract	10
	NaCl	5

	adjust pH to 7.0 with NaOH Agar	 10
NE	Glucose Yeast Extract 'Lab Lemco' powder (Meat extract) Casamino acids Adjust pH to 7.0 with NaOH Agar	10 2 2 2 20
Buffered NE	Use the same amounts as for 1l media, omit pH adjustment and make up to 900ml. Use 2ml buffer per 20ml plate.	
NMMP	$(NH_4)_2SO_4$ Casaminoacids $MgSO_4.7H_2O$ Trace elements solution Agar When plating add: NaH_2PO_4 / K_2HPO_4 buffer (0.1M, pH 6.8) Carbon Source (20%)	2 5 0.6 1ml 20 3ml per 20ml plate 0.5ml per 20ml plate
SFM	Soy Flour Mannitol Agar Use tap water in place of dH_2O	20 20 20
Buffered SFM	As SFM above, however make up to	

	900ml. Use 2ml buffer per 20ml plate.	
Buffered SF	Soy Flour Agar Make up to 800ml with dH_2O Carbon source (100mM) Buffer	20 20 2ml per 20ml plate 2ml per 20ml plate
Minimal Media	L-asparagine / $(NH_4)_2SO_4$ K_2HPO_4 $MgSO_4.7H_2O$ $FeSO_4.7H_2O$ Agar (Agarose may be used if a high purity medium is required) adjust to pH7.2 with NaOH Add filter sterilised carbon source to 1% w/v after autoclaving	0.5 / 1 0.5 0.2 0.01 10
Buffered Minimal medium	Components as above, minus pH adjustment. Make up to 800ml final volume, add 2ml buffer and 2ml carbon source per 20ml plate.	

2.4 *Commonly used solutions*

Solution	Components	Amount (g/l dH_2O unless otherwise stated)
10X TBE	Tris	108g
	Boric Acid	55g
	EDTA	9.3g
10X TBS	Tris	24.2
	NaCl	20
	Adjust pH to 7.5	
PBS	NaH_2PO_4	2.96
	Na_2HPO_4	11.5
	NaCl	5.84
	Adjust pH to 7.5	
Agarose gel (8%)	Agarose (high gel strength)	8
	1X TBE	1l
	Ethidium bromide solution (10mg/ml)	10µl
Bromophenol Blue DNA loading dye	Sucrose	40
	Bromophenol Blue	60 mg
	1 x TBE	10 ml
	dH_2O	90 ml
Buffer I	Tris	12.1
	NaCl	8.8

Buffer II / prehybridisation solution	20 x SSC 10% N-Lauryl sarcosine (w/v) 10% SDS (w/v) Blocking reagent Formamide dH_2O	25 ml 1 ml 200 µl 5 50 ml 24 ml
Buffer III	1 M Tris-HCl pH9.5 1 M NaCl 0.5 M MgCl2	100 ml 100 ml 100 ml
Denaturing Buffer	NaOH NaCl	20 87.75
Neutralisation Buffer	Tris NaCl pH 7.5	121 88
Antibody solution	Anti-Digoxigenin AP conjugate Buffer I	3µl 10ml
Colour solution	NBT/ BCIP Tablets (Roche) dH_2O	1 tablet 10ml
10% SDS	SDS	10
Stacking gel	Acrylamide dH_2O 1 M Tris pH6.8 10% SDS 10% TEMED 10% APS	650 µl 3 ml 1.25 ml 50 µl 5 µl 25 µl

Separating gel	Acrylamide	3 ml
(12%)	dH_2O	2.5 ml
	1.5 M Tris pH8.8	1.875 ml
	10% SDS	75 µl
	10% TEMED	7.5 µl
	10% APS	37.5 µl
Running buffer	Glycine	14.4 g
	Tris	4 g
	10% SDS	10ml
	dH_2O	up to 1 L
Transfer buffer	Tris	5.82
	Glycine	2.93
	SDS	0.035
	Methanol	200ml
TBS / Tween	10X TBS	100ml
	Tween 20	1ml
	Make up to 1l with dH_2O.	
Blocking buffer	TBS / Tween	1l
(5%)	'Marvel' Semi skimmed milk powder	5
Antibody solution	Blocking Buffer (5%)	20ml
for western blots	Antibody glycerol stock	1 - 5µl depending on conc. required
Coomassie stain	Coomassie Brilliant Blue R250	0.25
	Methanol	45ml
	dH_2O	45ml

	Glacial Acetic acid	10ml
Coomassie destain	Methanol	45ml
	dH_2O	45ml
	Glacial Acetic acid	10ml
Sample loading	1M Tris-HCl pH6.8	500µl
buffer	1M Dithiothreitol (DTT)	1 ml
	10% SDS	2 ml
	1% Bromophenol blue	1 ml
	Glycerol	1 ml
	dH_2O	4.5 ml
Sonication buffer	1M Tris / HCl, pH 8	2ml
	5M NaCl	1.6ml
	0.5M EDTA	1.2ml
	Protease inhibitor cocktail tablets (Roche)	1
	dH_2O	40ml
Column loading	Urea	23.27g
buffer 10mM	2M Imidazole	0.24ml
Imidazole	8X Phosphate buffer	6ml
	TritonX – 100	1ml
	dH_2O	to 48ml
Column washing	2M Imidazole	0.48ml
buffer 20mM	Other components as for loading buffer.	
Imidazole		
Elution buffer	2M Imidazole	7.2ml
300mM Imidazole	Other components as for loading buffer.	

FITC-coupled wheat germ agglutinin solution	Molecular probes	1 mg/ml in PBS, kept protected from light.
Fluo – WGA / PI solution	WGA solution PI solution 2% BSA in PBS	20 µl 4µl 10ml
PI wash solution	PI solution PBS	4µl 10ml
PI solution	Propidium Iodide dH_2O	25mg 1ml
Buffers for media additions	Individual acid solutions were made to desired concentration, then titrated to desired pH using NaOH.	

2.5 *Culture conditions*

E.coli cells were grown on LB media and incubated at 37°C either in a static temperature controlled incubator if being grown on plates, or shaken at 220rpm if liquid culture was used. Glycerol stocks of *E. coli* strains were kept at -70°C. Short term storage of cultures was possible by storing them at 4°C. *Streptomyces* cultures were incubated at 30°C on a variety of solid media. Spore suspensions were prepared for storage at -20°C by growing a lawn of the relevant strain on SFM and collecting the spores by suspending them in 2ml 20% glycerol. Once collected, the spores were filtered through sterile non – absorbent cotton wool by syringing to remove pieces of substrate mycelium.

2.6 *Antibiotic concentrations*

2.7 *DNA isolation, manipulation and analysis*

2.7.1 Ex traction of *Streptomyces* total DNA

Total DNA from *Streptomyces* was used for Southern blots and as a PCR template for amplification of specific genes. The method used to extract total DNA is based on the FastDNA® soil kit from Bio101. Breifly, cultures are grown on cellophane discs on agar plates to maximise yield of cells and minimise processing of pieces of agar. The cells are then harvested and transferred to a tube containing ceramic and silica lysis matrix. After addition of 978µl phosphate buffer and 122µl MT buffer (all buffers supplied with kit) the tubes were securely sealed and processed in the FastPrep® instrument for 30s at speed 5.5 to lyse the cells. After lysis, the tubes were spun at 13000rpm for 1min to pellet cell debris and lysis matrix. The supernatant was transferred to a fresh tube and 250µl PPS reagent was added before mixing by inversion. After a further centrifugation at 13000rpm for 1min, the supernatant was transferred to a 15ml tube and gently mixed with 1ml binding matrix. After settling, the matrix is transferred to a column and washed with 500µl SEWS – M solution before the DNA is eluted in 80µl dH_2O pre-warmed to 50°C.

2.7.2 Pol ymerase chain reaction (PCR)

PCR reactions were cycled on a PTC-200 DNA Engine (M.J. Research Inc.) using Finnzymes DynaZyme polymerase. Primers were designed with Beacon software from Premier Biosoft, and purchased from MWG. Primers used in this study can be

found in materials and methods. Reactions were carried out in thin walled tubes with a total reaction volume of 50µl. A typical reaction scheme is as follows:

DynaZyme buffer – 5µl

2.5mM DNTPs – 4µl

Sense primer (10pmol/µl) - 3µl

Antisense primer (10pmol/µl) - 3µl

Template DNA (approx 100µg/ml) - 2µl

dH_2O – 28.5µl

DMSO – 2.5µl

DynaZyme polymerase - 2µl

This reaction mix was used in the following program:

1. 93°C for 3min
2. 94°C for 30s
3. 63°C for 30s
4. 72°C for 3min
5. Repeat 2 – 4 30 times
6. 72°C for 8min
7. 4°C for ever

Primer name	Sequence
Forward	$^{5'}$ACAACGGCGAGGTGCTGGTCAGGAAGG$^{3'}$
Reverse	$^{5'}$CGGCAAGAAGAAGCAGGGCGGCAATCTC$^{3'}$

Table 2.3: Primers used in PCR reaction to amplify SAV4338.

2.7.3 D NA sequencing

Appropriate amounts of DNA were used (according to tables supplied with DTCS mix) and pre – heated at 96°C for 2min. This partially uncoils supercoiled DNA,

allowing better denaturation and binding in the reaction. The DNA was cooled on ice for 1min before the appropriate amount of primer was added from a 10pmol/µl stock (usually 2µl). This was then mixed with 8µl of DTCS quick start mix (Beckman Coulter) before the total reaction volume was made to 20µl using ddH_2O. The sequencing program most commonly used follows, however the number of cycles can be altered to increase / decrease the final DNA concentration: 1) 96°C for 20s, 2) 50°C for 20s, 3) 60°C for 4min, 4) 40x steps 1 – 3. After the reaction had finished, reactions were cleaned by ethanol precipitation of DNA, then separated using the Beckman CEQ 8000. Sequencing primers used are shown in table 2.4, however other primers used for site directed mutagenesis / PCR were also used.

Primer name	Sequence
1224	CGCCAGGGTTTTCCCAGTCACGAC
T7 universal primer	TAATACGACTCACTATAGGG

2.7.4 Souther n blotting

To confirm the successful insertion of a transposon into a piece of chromosomal DNA, it was necessary to perform a Southern blot. In brief, a restriction digest of chromosomal DNA from the mutant strain, chromosomal DNA from the parental strain and a positive control (often the transmitted plasmid / cosmid) was run on an agarose gel. The DNA was then transferred to a nitrocellulose membrane and detected immunologically using a specific probe. These steps are described in more detail below:

2.7.4.1 Creation of a probe labelled with digoxygenin

An appropriate sequence of DNA was chosen (often the transposon used in the original reaction) and excised from a gel after an appropriate restriction digest. Phage λ DNA digested *Hin*dIII was also obtained to hybridise with the marker. These samples were then labelled. To label the DNA, it was heated to 95°C for 5min then cooled on ice to denature it. A random priming mixture was then added to it (Hexanucleotide mix from Roche) along with the other components listed below. This reaction was then left overnight at 37°C. The reaction mix used is: 39 µl Denatured DNA, 5 µl Hexanucleotide mix, 5 µl dig-dNTP labelling mix, 1 µl Klenow polymerase. Before use, 10µl each of labelled probe and labelled marker DNA was transferred to 15µl prehybridisation solution, which will be called 'probe solution' from here.

2.7.4.2 Southern blot procedure

Fragments of digested chromosomal and control DNA were separated on an agarose gel until the loading dye had progressed to within 2cm of the end of the gel. The gel was then immersed in denaturing buffer for 15min with gentle shaking. This step was then repeated with a rinse in dH_2O between each immersion. After a further wash, the gel was immersed in neutralisation buffer for 20min with gentle shaking. This step was repeated, including a wash step as before. Concurrent with the final two immersions, a nitrocellulose membrane (Hybond-N, Amersham Pharmacia Biotech) was prepared and soaked in dH_2O for 20min and 10x SSC for 10min. DNA was then transferred from gel to membrane using pressure (90mm H_g) in a Stratagene Posiblot Pressure blotter. The membrane was supported on a piece of filter paper soaked in 10x SSC. After transfer, the membrane was baked for 1hr at 80°C.

2.7.4.3 Hybridisation

Following from the transfer of DNA from gel to membrane, the baked membrane was rolled into a cylinder shape with a piece of nylon mesh. This cylinder was put inside a hybridisation tube and wetted with dH_2O. The membrane was blocked by adding prehybridisation solution (approx 15ml) to the tube and incubating in a hybridisation oven at 42°C for 1hr. During this hour, the probe solution was boiled for 10min and then cooled on ice to ensure DNA denaturation. After the blocking step, the prehybidisation solution was discarded and the probe was added. The membrane was incubated at 42°C in the prehybridisation oven overnight. The following day, the probe was collected and stored at -20°C for future use. The membrane was washed twice at 42°C for 15min in wash solution 1, then it was washed twice in wash solution 2 at 68°C.

2.7.4.4 Immunological detection

After a brief wash in buffer 1 at room temperature, the membrane was blocked again in prehybridisation solution at room temperature for 30min. Following another brief wash in buffer 1, 15ml of antibody solution was added to the tube and incubated for 30min. Following the hybridisation, two final 15min washes in buffer 1 were carried out. To visualise the hybridisation, the membrane was sealed in a plastic bag with 10ml colour solution and incubated at 37°C for 30min – 2hr depending on the amount of DNA present and the efficiency of the hybridisation. After sufficient band intensity had been reached, the reaction was stopped by washing in tap water.

2.7.5 Isolation of plasmid / cosmid DNA

This was performed using the 'Wizard' miniprep kit from Promega. A culture of *E. coli* containing the plasmid / cosmid was grown in LB broth with appropriate

antibiotic selection for 6hr or overnight. The cells were then pelleted using a bench top centrifuge, the supernatant was discarded and the pellet was resuspended in 250µl resuspension solution. 250µl lysis solution was then added and the mixture gently inverted to lyse the cells and release DNA into solution. The proteins and other cellular components were precipitated by addition of neutralisation solution and a forceful mix by inversion, and they were then removed by spinning at 13000rpm for 10min followed by careful decanting of the supernatant to a spin column. The column contains a matrix that binds cosmid / plasmid DNA, and the supernatant was passed over this by spinning at 13000rpm for 1min. The flow – through was discarded and the column was then washed twice, first in 750µl then in 350µl wash solution using centrifugation as before. Finally, DNA was eluted using 80µl ddH_2O pre – heated to 50°C.

2.7.6 D NA manipulation using enzymes

Restriction digests were performed using restriction enzymes purchased from New England Biolabs (NEB). All reactions were carried out in accordance with the manufacturers' instructions. T4 DNA ligase was purchased from Invitrogen and used in accordance with the manufacturers' instructions. Calf intestinal phosphatase (CIP) was purchased from NEB and reactions were performed as instructed.

2.7.7 Visuali sation of DNA using agarose gel electrophoresis

To see the result of a particular DNA manipulation (e.g. a restriction digest), DNA fragments were separated using agarose gel electrophoresis. Briefly, an agarose gel was made to a final strength of 0.8% w/v containing a final concentration of 0.1µg/ml ethidium bromide. DNA was mixed in a ratio of 1:5 loading buffer to DNA, and then the gel was immersed in a tank containing 1X TBE running buffer. DNA was loaded

into wells using a pipette and allowed to settle at the bottom of each well. After the addition of an appropriate marker (λ *Hin*dIII in most cases) the DNA was forced through the gel by passing an electric current through the gel, usually the voltage was kept constant at 100V, although for high current runs the current was kept constant to avoid fluctuations due to heating. Once the DNA had migrated the desired distance, the bands were visualised by use of a 254nm transilluminator.

2.7.8 Isolation and purification of digested DNA fragments

If one particular fragment is required, it can be excised from a gel and purified using the GFX kit from GE healthcare (formerly Amersham Biosciences). DNA in solution can be purified using ethanol precipitation or using the PCR purification kit from Quiagen. Both the kits were used according to the manufacturers' instructions the principal being similar for both kits. Firstly, the DNA is mixed with a binding buffer and loaded onto a matrix using centrifugation. In the case of DNA from gels the agarose must be melted at 60°C in the presence of the buffer. After loading, the matrix is washed and the DNA is then eluted in ddH_2O pre – warmed to 50°C.

2.7.9 Site directe d mutagenesis

This technique is used to introduce a specific change in a region of DNA. This could be to change the identity of an amino acid in the protein translated from a particular gene, or it could be (as in this case) to introduce a restriction site at a convenient point. The QuikChangeXL® site directed mutagenesis kit from Stratagene was used to introduce a new restriction site at the start of SCO3848, and it was also used to create a new site in the DNA coding for the transmembrane region of the expressed protein. Primers were designed which contained the new bases and flanking sequences either side (see below).

Name	Sequence
MidKin1	CGATCTTCCTGCATATGGCGGGCGTCC
MidKin2	GGACGCCCGCCATATGCAGGAAGATCG

The reaction mixture was as follows:

5 µl	10 x reaction buffer
10 ng	dsDNA template
125 ng	primer 1
125 ng	primer 2
1 µl	dNTP mix
3 µl	QuikSolution
1 µl	PfuTurbo DNA polymerase (2.5U/µl)
up to 50 µl	ddH_2O

Following a mixing step, the reaction was run on a thermal cycler using the following program:

1) 95°C	1 min
2) 95°C	50 sec
3) 60°C	50 sec
4) 68°C	1 min per Kb plasmid length
5) Goto 2 17 times	
6) 68°C	7 min

After the thermal cycling reaction had finished, the reaction was incubated at 37°C for 1hr with 1µl *Dpn*I to digest any remaining parental DNA. DNA was then transformed into *E. coli* XL10-Gold cells supplied and grown on LB agar plates containing the appropriate antibiotic at 37°C overnight. Colonies were selected, plasmid DNA

extracted, and confirmation of the correct mutation was obtained using restriction digests and sequencing.

2.7.10 Transposon mutagenesis

This technique was used to introduce a transposon into a gene *in vitro* before transforming it into its parent organism for examination of the *in vivo* effects of the disruption. This is based on the EZ Tn kit from Epicentre, however the transposons used are in – house designs based on Tn5. The reaction mix used is as follows:

10x buffer	1µl
Target DNA	200ng
Transposon DNA	13ng
Transposase	1µl
ddH_2O	to 10µl

After incubation at 37°C for 2hr, 1µl stop solution is added and the reaction is incubated at 70°C for 10min. The reaction is then transformed into *E. coli* using standard techniques.

2.7.11 Transformation of competent *E. coli*

2.7.11.1 Production of electrocompetent *E. coli*

A 1/100 dilution of a fresh overnight *E. coli* JM109 / ET12567 culture was inoculated into 500ml LB broth. At an OD_{600} 0.5 – 0.7 the cells were chilled on ice for 20 min then harvested by centrifugation at 4000rpm for 15min at 4°C. In all following procedures the cells were kept on ice as much as possible, and all glycerol solutions

used were ice cold. The supernatant was decanted and the pelleted cells carefully resuspended in 500ml 10% glycerol. The cells were then pelleted and resuspended in 250ml 10% glycerol. Finally cells were pelleted, resuspended in 20ml 10% glycerol, pelleted again and resuspended into 2ml 10% glycerol. The 2ml suspension of electro-competent JM109 cells was aliquoted out and stored at –70˚C.

2.7.11.2 Production of heat – shock *E. coli*

An *E. coli* JM109 / ET12567 culture was prepared as for electrocompetent cells, however after the initial pelleting stage the cells were resuspended in one half volume of ice cold, sterile 50mM $CaCl_2$ and incubated on ice for 20min. The cells were then pelleted again and resuspended in one tenth volume of 50mM $CaCl_2$. Sterile glycerol was added to a final concentration of 20% w/v, cells were dispensed into 100µl aliquots and stored at -70ºC.

2.7.11.3 Transformation of competent cells

Electrocompetent cell aliquots were thawed on ice before DNA was added and mixed by gentle pipetting. This mixture was then transferred to an ice cold electroporation cuvette. This was then placed in a MicroPulser™ (BioRad) and electroporated using program EC1. 400µl LB broth was then added and the culture was shaken at 225rpm and 37ºC for 1hr before being plated on LB agar plates containing the appropriate antibiotic selection.

Heat shock competent cells were thawed on ice before DNA was added, and after mixing, the cells were left on ice for 1hr. The cells were then heat shocked for 2min at 42ºC. After this treatment, the mixture was briefly cooled on ice before 400µl LB

broth was added. From here on, the procedure is identical to the electroporation procedure.

2.7.12 Genetic engineering of *Streptomyces*

2.7.12.1 Interspecific conjugation

ET cells were transformed with a plasmid / cosmid containing *oriT* and antibiotic selection other than kanamycin or chloramphenicol (used to maintain *dam-* and pUZ8002). After they had grown on the LB agar plate, several colonies were selected and grown together in LB broth containing the appropriate antibiotic selection for 4 – 6hr. Following this stage, the cells were pelleted and simultaneously 20µl of a dense *Streptomyces* spore suspension was placed in 500µl 2xYT broth and incubated for 10min. During the incubation, the ET cells were gently washed twice with LB broth to remove residual antibiotic. After the spores had been treated and cooled to room temperature, the two species were mixed and suspended together. The Suspension was pelleted using centrifugation then most of the supernatant was decanted. The cells and spores were suspended in the residual broth and plated on SFM containing 10mM $MgCl_2$ and incubated at 30°C overnight. The following day, 0.5mg nalidixic acid and antibiotic of the appropriate concentration was overlayed on each 20ml plate (nalidixic acid kills *E. coli* while *Streptomyces* are resistant to it).

2.7.12.2 Selection of *Streptomyces* double crossovers

After 2 – 3 days incubation, single colonies appear on the overlayed SFM plates. These colonies can then be screened for successful integration in the *Streptomyces* chromosome by double crossover. Two sets of plates are prepared, one containing the appropriate concentration of the antibiotic whose resistance gene is carried on the

plasmid / cosmid (e.g. kanamycin), the other containing the antibiotic whose resistance gene is carried on the transposon (e.g. apramycin). Replica plating is then carried out for a large sample of the individual colonies (100 – 200). Any colonies which grow on apramycin but not kanamycin plates are likely to be colonies containing a double crossover, and this can be confirmed using Southern blotting.

2.8 *Protein overexpression, isolation and analysis*

2.8.1 O verexpression of His – tagged protein in *E. coli*

BL21(DE3) or Novablue *E. coli* cells were transformed with the appropriate plasmid (in this study a pET16k derivative) and grown on LB agar plates containing kanamycin. A single colony was picked and grown overnight at 37°C, 225rpm in liquid LB plus kanamycin. The overnight culture was used to make a high cell density stock solution by pelleting and adding 50% glycerol. This stock was used for future inoculations. 1ml of the overnight culture was inoculated into 24ml LB plus kanamycin and grown at 37°C, 225rpm in a shake flask for 4 - 6hr. To induce overexpression, IPTG was added to the culture to a final concentration of 10mM and shaken at 30°C, 225rpm for 1 – 3hr. Following induction, the cells were pelleted and resuspended in 5ml sonication buffer. To lyse the cells and fragment the DNA contained within whilst maintaining intact proteins, sonication was used to disrupt the cells. Treatments varied according to density of cell suspension and power of sonicator used, however all treatments were continued until the lysate would not continue to clear with additional treatment. The lysate was either then used to give whole cell extract data, or used in further steps.

2.8.2 Ex traction of membrane fraction

The clarified lysate was spun in a centrifuge at 13000rpm for 5min to remove intact cells and inclusion bodies. This is the insoluble fraction. The supernatant (containing membrane and soluble fractions) was then taken and centrifuged at 450000rpm for 1hr in an ultracentrifuge. The supernatant (soluble fraction) was then decanted and kept, whilst the membrane fraction was suspended in sonication buffer containing 2% TritonX – 100. Occasionally, a brief sonication treatment was necessary to totally dissociate the membranes.

2.8.3 Pu rification of expressed protein

Before purification, the cell pellet produced from induction must be resuspended in binding buffer before sonication to cleared lysate as before. The technique used to purify protein was the Hi – Trap system from Amersham Biosciences. The system relies on the affinity of poly – his tagged proteins for Ni^{2+} ions and the dissociative properties of imidazole. The system comprises a column which must be pre – loaded with Ni^{2+}, and a syringe. During all steps, it is vital that air does not enter the column, therefore all solutions were added to the column 'drop to drop' to avoid bubbles. Bubbles were also removed from the syringe before joining it to the column. After loading and washing excess Ni^{2+} from the column with 5ml binding buff er, the sample is loaded using the syringe and passed over the column. The flowthrough from each run is collected for analysis. After this step, the column is washed with 10ml wash buffer. The next buffer used can have differing concentrations of Imidazole, depending on the affinity of the target protein for the nickel ions. The elution uses 5ml buffer of the appropriate imidazole concentration and fractions are collected as before. The fractions can be observed on a Coomassie stained gel.

2.8.4 Isolation of phosphoproteome

For this procedure, the PhosphoProtein purification kit from Qiagen was used. This kit uses affinity chromatography to bind proteins containing phosphorylated amino acids to the column. This allowed non – phosphorylated proteins to wash through the column to be saved as the non – phosphorylated fraction. Once the bound, phosphorylated proteins had been washed, the phosphoproteome could be eluted from the column and then further concentrated using Nanosep columns (supplied with the PhosphoProtein purification kit). The buffers did not deviate from the conditions stipulated in the product manual, however the lysis buffer was used in a sonication step in order to disrupt the cell walls of the bacterial cells. Sonication was performed as in section 2.8.1.

2.8.5 Pol yacrylamide gels

Polyacrylamide gels can vary in size and strength according to the resolving power required. Gels were made according to the formulae above, with variation in polyacrylamide concentration as necessary. Protein samples were mixed with loading buffer and boiled for 5 min to maximise dissociation. The gels were run at a constant voltage of 90V in running buffer.

2.8.6 Sta ining

2.8.6.1 Coomassie staining

Polyacrylamide gels were incubated with Coomassie stain for at room temperature for 1hr with gentle agitation. After recovery of the stain for re-use, the gel was destained in destain solution in several 20min steps, renewing the destain solution at each step.

The steps were repeated until protein bands were visible against background staining. Resolution was improved by leaving gels in water overnight.

2.8.6.2 Silver staining

All solutions were made using supplied components according to the manufacturer's instructions. All steps were performed at room temperature and with gentle agitation. After the gel had been run, it was placed in fixing solution for 30min. Next the gel was washed 3 times for 5min each in dH_2O. Whilst the washes were being performed, 0.25ml 25% glutaraldehyde was added to the sensitizing solution. After the final wash, sensitizing solution was added followed by incubation for 30min. A further series of 3 x 5 minute dH_2O washes were performed, and during this step 20µl of 37% formaldehyde was added to the silver reaction solution. Following the washes, the silver reaction solution was added to the gel and incubated for 20min, followed by two 1min dH_2O washes. During these washes, 10µl 37% formaldehyde was added to the developing solution. This solution was used to develop the gel. Developing time varies between gels depending on band intensity, typically between 1 and 5 minutes is required. To stop the reaction, remove the developing solution and add stopping solution. Leave for 10min before recording any pictures.

2.8.7 Wes tern blotting

During the period in which a gel is running, preparation of the blotting PVDF membrane can be achieved. This requires activation of the membrane in pure methanol for 5min, followed by a brief dH_2O wash. The membrane is then saturated in transfer buffer. After the gel is prepared, it is also saturated in transfer buffer along with two pieces of thick blotting paper cut to the size of the gel with approximately 0.25cm overhang at each edge. The transfer cassette is made up by first laying a piece

of blotting paper on the anode. This is pressed onto the surface using a glass rod to expel air bubbles, and air bubbles are removed from every subsequent step in a similar manner. The membrane is added next, followed by the gel, and finally the second piece of filter paper is placed on top of the gel. The cassette is run at 20V for 20min, increasing by 5V and 5min for every extra minigel. After transfer, the membrane is washed in TBS / Tween briefly before being blocked in 10% milk powder in TBS / Tween for 1hr at room temperature. Following the blocking step, the membrane must be thoroughly washed in TBS / Tween, at least 4 x 5min. The membrane is then incubated with primary antibody, which can either be done for 1 – 2hr at room temperature or overnight at 4°C. If a secondary antibody is required, it can be added after the membrane has been washed. After the antibody stages, the membrane is again washed and can be visualised using photographic film and colour solution. The system used was the ECL system from Amersham Biosciences.

2.8.8 M ass spectrometry

2.8.8.1 Destain of Coomassie stained gel pieces

First, protein was purified as much as possible using the method described above. Then, a 1D polyacrylamide gel was run, stained with Coomassie blue, and the band in question was excised from the gel using a scalpel and transferred to a clean eppendorf tube. The pH was corrected with a brief wash in 100mM NH_4HCO_3. The gel pieces were then destained with the addition of 200µl 50% acetonitrile / 50% 50mM NH_4HCO_3 and vortexed, followed by a 10min incubation at room temperature. This step was repeated until all stain had been washed from the gel pieces. Excess lquid was removed, and the pieces were washed for 5min in 200µl 50% acetonitrile / 50% 50mM NH_4HCO_3. The gel pieces were dehydrated with a 5min incubation with 100%

acetonitrile, followed by air drying for 10min at room temperature. Gel pieces were then washed with 100µl 100mM NH_4HCO_3 for 5min with vortexing, then twice with 50% acetonitrile / 50% 50mM NH_4HCO_3, 5min each step with vortexing. Dehydration with 100µl acetonitrile was then performed as before. Following removal of excess liquid (speedvac if necessary), the gel pieces were shrunken and white.

2.8.8.2 Trypsin digestion

Approximately 20µl of trypsin solution was added to the gel pieces (i.e. enough to completely cover them). The gel pieces were then rehydrated in buffer containing 50mM NH_4HCO_3 and trypsin at 4°C for 30min. The low temperature aids the absorption of trypsin into the gel with minimal autoproteolysis. After a brief spin, enough NH_4HCO_3 to cover the gel was added. The digest was then incubated overnight at 37°C.

2.8.8.3 Peptide extraction

Any supernatant present was transferred to a clean eppendorf tube, and 30µl (enough to cover) of 50% acetonitrile / 2% formic acid was added to the gel pieces. After a 20min incubation with vortexing, the pieces were sonicated briefly, then spun and the supernatant was added to the original supernatant. This step was repeated to give a final volume of around 60µl. This volume was vortexed, then reduced in volume to 5 - 10µl in the speedvac. After a final 10min centrifugation at 13000rpm, the sample was ready for loading onto the LC-MS/MS. Samples can be stored in the freezer.

2.9 *Gene expression analysis using luciferase reporter gene*

96 well culture plates (Greiner bio-one) were used to quantify luxAB expression assays. 300 µl agar was pipetted into each well and 2 µl spore suspension was

aliquoted into each well. Cultures were then incubated at 30°C. At appropriate time intervals plates were exposed to N-decanal; approximately 1 ml N-decanal was pipetted onto filter paper cut to the size of the plate. The saturated paper was laid on top of the plate for 1 min. This allowed the luciferase enzyme to undertake the oxidation of N-Decanal. The resulting luminescence (arbitary units) was measured with Lucy1 Luminometer (LabTech). On each plate, samples were measured in triplicate.

2.10 *Bioinformatic techniques*

2.10.1 Sequence selection and manipulation

Sequences were selected by performing structural homology searches using BLASTP, CDART and SMART (Altschul et al. 1997; Letunic et al. 2004). The amino acid sequences of PASTA domains from *S. coelicolor* were used as a query sequence. A representative sample of organisms was chosen from the returned hits, and the amino acid sequences were checked for the correct structure using SMART. For each sequence, the corresponding DNA sequence was downloaded from the NCBI database and saved in FASTA format. Using the SMART structural prediction, it was possible to use a translation of the whole DNA sequences to select in – frame portions of the DNA sequence that matched the amino acids contained within each domain (for example, in the case of *S. coelicolor*, one sequence corresponding to the STPK domain, one for the whole extracellular domain starting from the first PASTA domain, and one for each PASTA domain (four in total) were selected). All sequences were saved in FASTA format as plain text files (.txt).

2.10.2 Alignment of sequences

Once the sequences had been converted into FASTA format, it was necessary to align the DNA sequences according to their amino acid sequences. This allows individual codons to be compared against each other. The alignments were achieved using MEGA 3.1 (Kumar, Tamura, and Nei 2004), which can translate multiple DNA sequences into the corresponding amino acid sequences. Using CLUSTALW, the amino acid sequences were aligned (leaving gap penalty settings at default values). Following the alignment of the amino acids, the corresponding DNA alignment was selected and saved for future use as a FASTA (.fas) file.

2.10.3 Calculation of synonymous and nonsynonymous substitution rates

Synonymous codon usage (K_s) and non-synonymous codon usage (K_a) was calculated using the Nei – Gojobori method, including the Jukes – Cantor correction for multiple substitutions at the same site. The following is a brief outline of the formulae used in these calculations:

Number of Sites (S or A)

The numbers of potential synonymous and nonsynonymous sites can be computed. For each pair of sequences, the average number of synonymous or nonsynonymous sites is reported.

Number of differences (N_s or N_a)

These are simple counts of the number of synonymous (N_s) and nonsynonymous (N_a) differences. To compare these two numbers, you must use the p-distance because the number of potential synonymous sites is much smaller than the number of nonsynonymous sites.

p-distance (p_s or p_a)

The count of the number of synonymous differences (p_s) is normalized using the possible number of synonymous sites (S). A similar computation can be made for nonsynonymous differences.

Jukes-Cantor correction (K_s or K_a)

The p-distances computed above can be corrected to account for multiple substitutions at the same site.

Formulae:

$p_s = S/N_s$

$p_a = A/N_a$

$K_s = -{}^3/_4 \ln(1-{}^4/_3 p_s)$

$K_a = -{}^3/_4 \ln(1-{}^4/_3 p_a)$

Formulae for the variances of each of the above can be found within MEGA3.1 documentation. As well as MEGA3.1, DNASP (Rozas and Rozas 1999) was used to compute K_a/K_s in sliding window analysis.

2.10.4 Sliding window analysis

Sliding window analysis is used to compare aligned sequences a few codons at a time. Thus, if a sequence has a very low K_a/K_s value overall, sliding window analysis will allow areas under weaker selective pressure to be observed without being masked. Different window and step sizes can be chosen, depending on the degree of detail needed in a particular analysis. When choosing step and window sizes, make sure each is a multiple of 3. This ensures codons are being compared in – frame.

2.10.5 Construction of phylogenetic trees

Phylogenetic trees were constructed using PHYML (Guindon and Gascuel 2003). This program constructs a tree using maximum – likelihood methods, and the trees in this work were all constructed according to the Tamura – Nei model of nucleotide substitution with 500 bootstrap replicates.

2.11 *Microscopic techniques*

2.11.1 Preparation of samples for staining

Samples were prepared for FITC/WGA staining using the inserted coverslip method. A coverslip is inserted into solid agar media at an angle of approximately 30° from the horizontal. The inoculum is then pipetted into the acute angle between the agar and coverslip, and then incubated for the required period of time. After incubation, the coverslip is removed from the agar and the cells on the coverslip are fixed in either fixing solution or pure methanol prior to staining. For atomic force microscopy (AFM) the sample is not fixed and it is viewed directly with no further preparation. For Baclight staining, the recommended concentrations of PI (1µl/ml) and RedoxSensor (1µl/ml) stain are added to 20% glycerol solution, then the coverslip is inverted and the cells sit in the solution.

2.11.2 Staining of samples and preparation of slides

After fixation, the sample is rehydrated in PBS for 5min. An optional treatment with lysozyme solution can be carried out for 1min at this point, followed by a PBS wash. 2% BSA/PBS is then added for 5min, after removal the staining solution (WGA-FITC / PI in 2% BSA) is added and incubated at room temperature for 20 – 25min. The stain is then removed and the sample is washed at least eight times in PBS / PI. Slides

are prepared by floating the coverslip (cells on underside) on a small quantity of glycerol. Any excess glycerol is removed and the edge of the coverslip is sealed with nail varnish.

2.11.3 Microscopy

Microscopy was performed with an epifluorescent microscope using appropriate filters for the stains used. Digital images were captured using a camera mounted on the microscope.

2.12 *Materials and Methods References*

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25**:3389-3402.

Flett, F., V. Mersinias, and C. P. Smith. 1997. High efficiency intergeneric conjugal transfer of plasmid DNA from Escherichia coli to methyl DNA-restricting streptomycetes. FEMS Microbiol Lett **155**:223-229.

Guindon, S., and O. Gascuel. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol **52**:696-704.

Kumar, S., K. Tamura, and M. Nei. 2004. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief Bioinform **5**:150-163.

Letunic, I., R. R. Copley, S. Schmidt, F. D. Ciccarelli, T. Doerks, J. Schultz, C. P. Ponting, and P. Bork. 2004. SMART 4.0: towards genomic data integration. Nucleic Acids Res **32 Database issue**:D142-144.

Rozas, J., and R. Rozas. 1999. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis
10.1093/bioinformatics/15.2.174. Bioinformatics **15**:174-175.

3 Phylogenetic analysis of *pknB* orthologues 1: Evolution of PknB – like kinase genes.

3.1 Introduction

STPK's were first discovered in bacteria in the early 1990's from *Myxococcus xanthus* (Munoz-Dorado, Inouye, and Inouye 1991). As more genomes were published, it became clear that STPK's were widespread in bacteria and archea (Leonard, Aravind, and Koonin 1998) and therefore were of ancient origin. Further studies showed similar G+C content for STPK genes and their host genome (Ogawara et al. 1999; Han and Zhang 2001). The evolution of the STPK's within bacteria has received less attention than characterisation of individual genes. This functional characterisation has revealed that STPK's are part of the signalling machinery for a very diverse set of processes. To take an example, one *in silico* study of *S. coelicolor* has revealed 34 putative STPK genes with predicted functions including cell division, energy metabolism, DNA repair and many more (Petrickova and Petricek 2003).

Many STPKs have additional domains, as is the case with PknB, and these also have diverse functions. PknB contains four 'PASTA' (Penicillin binding protein And Serine / Threonine kinase Associated) domains, predicted to interact with peptidoglycan monomers (Yeats, Finn, and Bateman 2002). BLASTP and CDART searches using *S. coelicolor* PASTA domains as a query sequence revealed a family of transmembrane STPKs exclusively in Gram – positive organisms, exclusive to the firmicutes and actinobacteria.

3.2 *Inference of function from genetic context*

The context in which a gene is found can often reveal clues about its function. Genes that are clustered together can be predicted to code for proteins that interact with each other (Wolf et al. 2001). This only holds true if the same gene order is conserved across different species. The conservation of gene order can be less than obvious, often with genes being rearranged into different genetic contexts as the process of vertical transfer happens over evolutionary time (Lathe, Snel, and Bork 2000). This continuing process of rearrangement can at first glance appear to be caused by lateral transfer of genes between species – for example the presence of cell division genes in the midst of a cluster dedicated to DNA replication and ribosome assembly described below. In order to successfully identify a horizontally transferred gene, one must examine more than just the context in which it is found, and in the cases below codon usage and G+C content will be used to identify any anomalous sequences.

3.2.1 Pk nB gene cluster

pknB and its orthologues lie within a highly conserved gene cluster that is present in all sequenced actinomycete genomes and lies close to the origin of replication in all the organisms studied, including two that have undergone dramatic reductive evoloution (*Tropheryma whipplei* and *Mycobacterium leprae*). This cluster has a high proportion of conserved genes, with some variation in gene content between organisms (see figure 3.1). The genes in the cluster can be divided into two groups, based on prediction of function. The first group is involved in cell wall modification (*pbp, rodA, crgA*) and the second group is involved in signalling (*fha, ppc2c, pknB*). The wider context in which this cluster is found in *S. coelicolor* points superficially to a horizontal transfer event. There are two tRNA genes present, one at either end of the

cluster, and this is often indicative of DNA rearrangement (Blum et al. 1994; Koonin, Makarova, and Aravind 2001; Chen 2006). In the case of this cluster, codon usage statistics for each gene within it show very similar patterns as for the genome as a whole (fig 3.2). The story is the same for G+C content at each position, with no significant differences at each codon position (table 3.1). This conformity of DNA composition is also reflected in the phylogeny of each gene within the cluster, with trees drawn for each of the conserved genes within the cluster mirroring the phylogeny of accepted molecular clocks such as *gyrA* and 16S rDNA genes (fig 3.3).

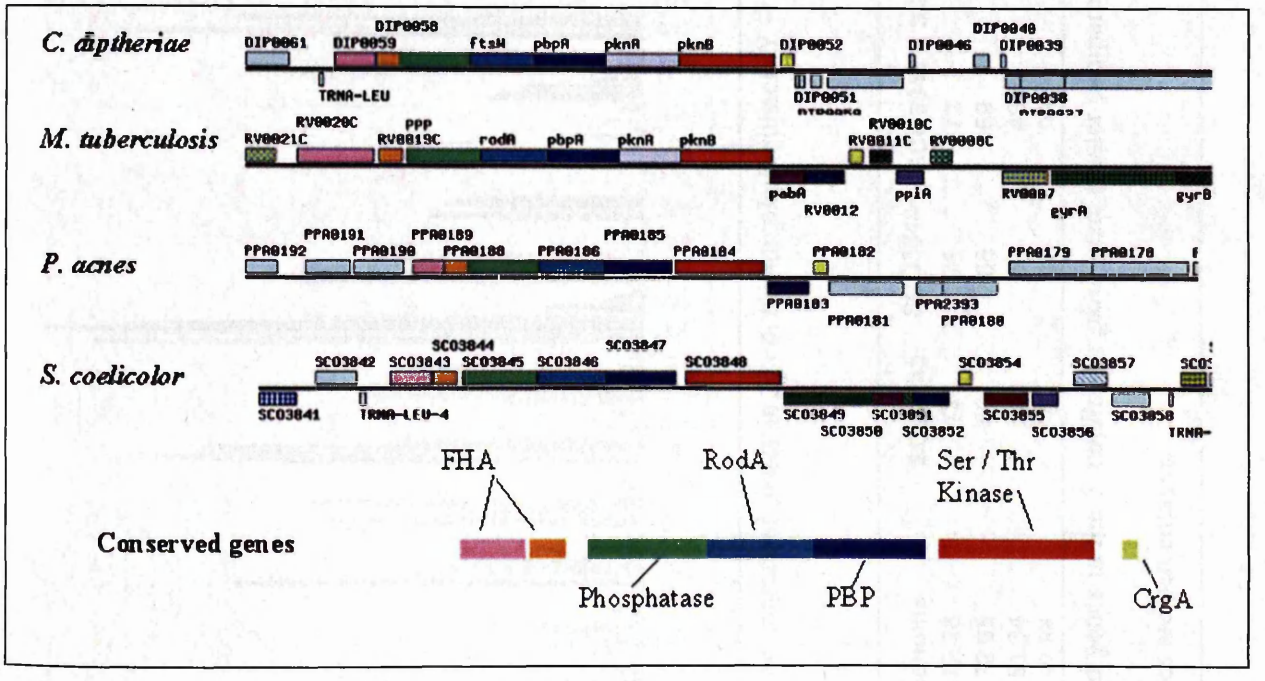


Figure 3.1: Examples of actinobacterial *pknB* gene clusters. Conserved genes are shown below the examples.

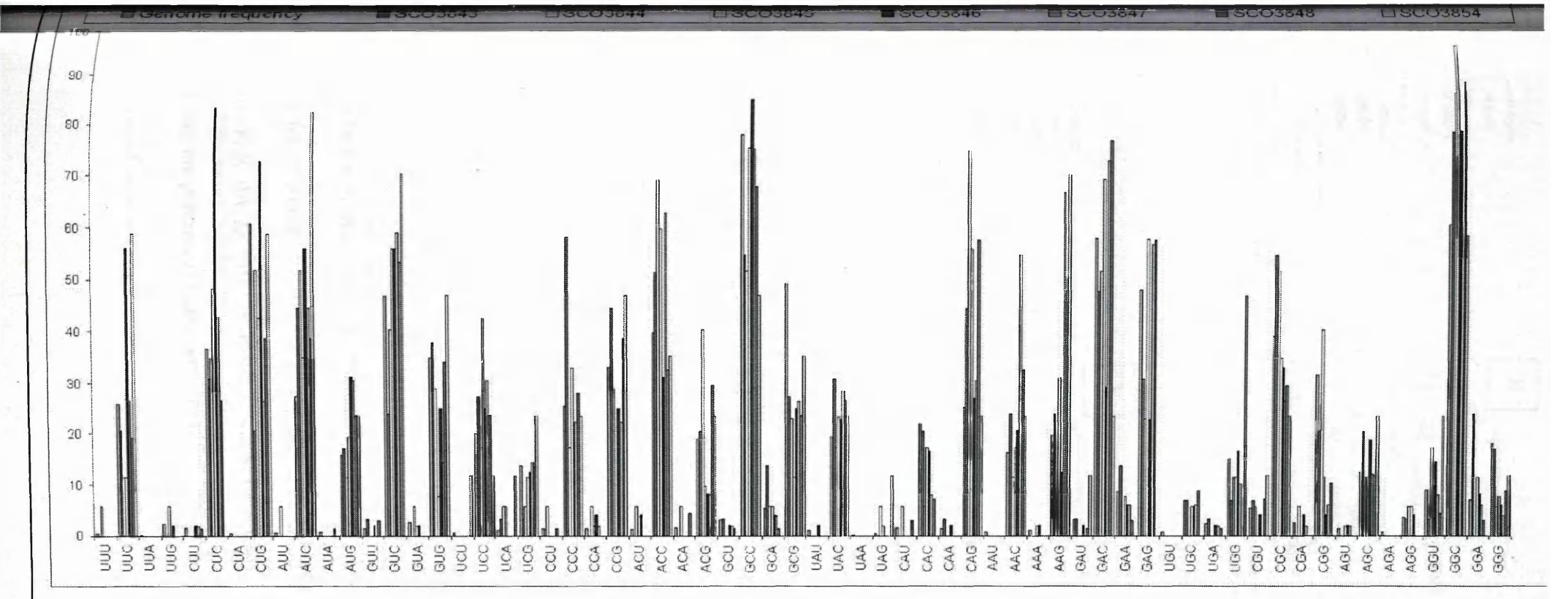


Figure 3.2: Bar graph showing codon usage for conserved genes in *S. coelicolor pknB* cluster. Y – axis shows frequency of use per thousand codons, X – axis shows RNA codons.

	Genome	SCO3843	SCO3844	SCO3845	SCO3846	SCO3847	SCO3848	SCO3854
Total G+C	72.28	72	71	71	68	66	69	64
1st	72.65	69	68	69	67	60	68	52
2nd	51.34	56	63	49	55	43	45	47
3rd	92.84	92	92	96	93	97	96	93

Table 3.1: G+C content of conserved genes in the *S. coelicolor pknB* gene cluster compared to the genome as a whole. Also shown is G+C content at each codon position. Values are percentages.

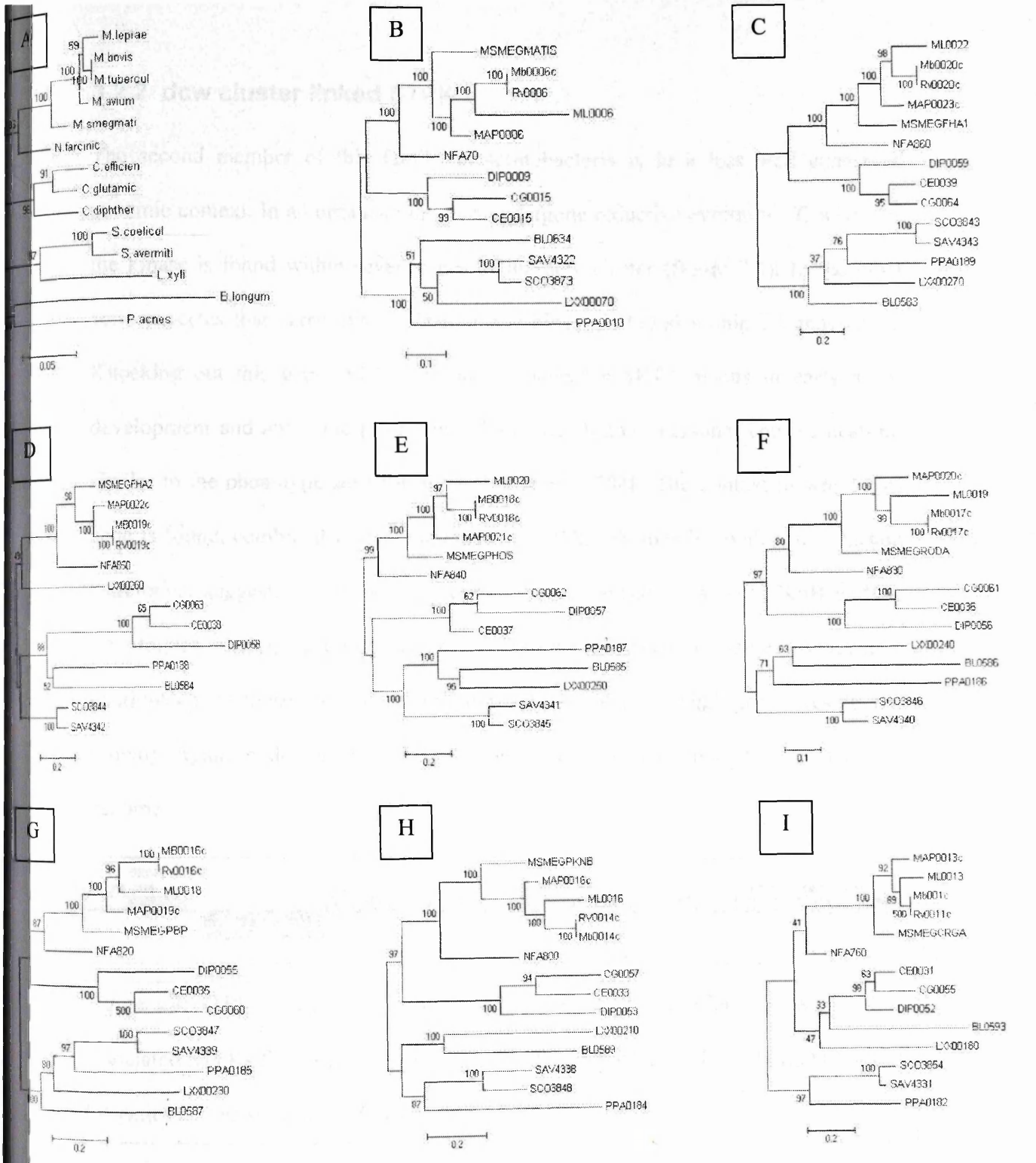


Figure 3.3: Maximum likelihood trees drawn for conserved genes in the *pknB* gene cluster, and two molecular clock genes for reference. A: 16S rDNA; B: *gyrA*; C: SCO3843 orthologues; D: SCO3844 orthologues; E: Phosphatase; F: RodA; G: PBP; H: PknB; I: CrgA. Each tree was drawn using PHYML and bootstrap values shown at nodes are percentages calculated from 500 replicates.

3.2.2 dcw cluster linked STPK

The second member of this family in actinobacteria is in a less well conserved genomic context. In an organism that has undergone reductive evolution, *T. whipplei*, the kinase is found within seven genes of the dcw cluster (figure 3.4). In the other actinomycetes that carry this duplication the kinase is found within 20 genes of it. Knocking out this gene (SCO2110) in *S. coelicolor* M600 results in early aerial development and antibiotic production (Jung and Buttner, personal communication) similar to the phenotype seen for disruption of SCO3848. The context in which this gene is found, combined with the interaction of PASTA domains with beta – lactam antibiotics, suggests a role for this kinase in cell division. As with PknB and its orthologues, phylogenetic trees drawn for this kinase indicate the same phylogenetic relationships as those found between the accepted molecular clock genes (results not shown). Again, codon usage and G+C content is similar to that of the rest of the genome.

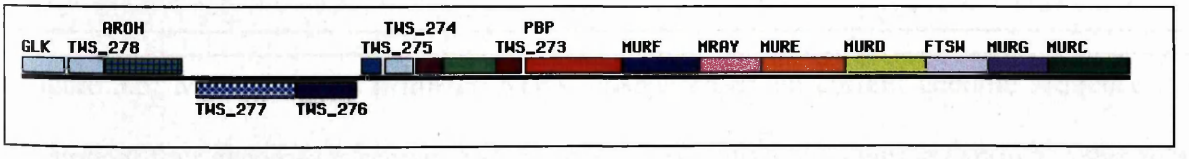


Figure 3.4: STPK linkage in *T. whipplei*. TWS_277 is the equivalent of TW555 (dcw associated STPKs from strains Twist and TW/08 respectively). PBP, MUR, MRA and Fts genes are the start of the dcw cluster.

3.2.3 Firmicute kinase linkage

The firmicutes generally contain a single STPK gene which associates with a phosphatase gene. The genetic locus of these genes is well conserved, however no obvious functional linkage can be inferred from it as there appear to be no genes

dedicated to cell wall processes nearby (figure 3.5). Recent genome sequencing projects have revealed linkage between the 'actinobacterial' cluster and *priA* in some firmicutes (see next section). It has been shown that mutation of the STPK (*prkC*) affects spore formation, increases cell number and disrupts biofilm formation in *Bacillus subtilis*, with mutation of the phosphatase (*prpC*) having an opposite effect (Madec et al. 2002; Madec et al. 2003). The orthologue in *Streptococcus pneumoniae* (*stkP*) has been implicated in virulence (Echenique et al. 2004). As for the other members of this kinase family, G+C content and codon usage are the same as for the host genome, and the phylogeny is the same as for the molecular clock genes. (results not shown).

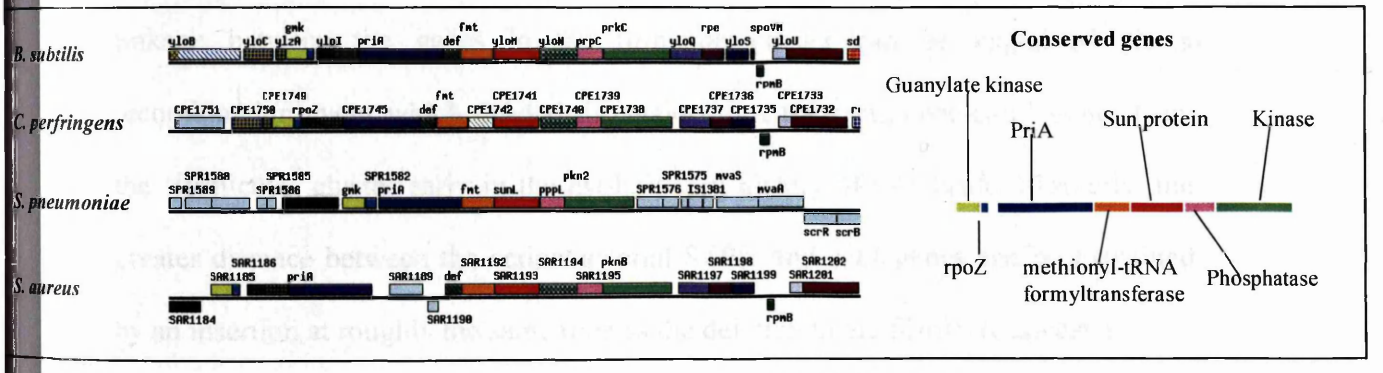


Figure 3.5: Most common firmicute STPK linkage based on current genome sequences. A sample of four genomes is shown, and the products of conserved genes are shown to the right.

3.3 Evidence for linkage of firmicute and actinobacterial clusters

The linkage between STPK genes and cell wall modification genes is clear in the actinobacteria, however until recently there has been little evidence for such a role in the firmicutes. The experimental evidence for the STPK function in firmicutes provides a tantalizing suggestion of cell – wall related activity, however this is not

supported by the genetic linkage shown in the majority of sequenced genomes (figure 3.5). For this evidence we must look at three recently sequenced species of bacteria, *Desulfitobacterium hafniense, Moorella thermoacetica* and *Desulfotomaculum reducens.* These bacteria belong to the *Clostridia* family, and so are members of the firmicutes (NCBI taxonomy), In these bacteria there is a cluster resembling the one found in actinobacteria coding for a single transmembrane PASTA domain STPK, a PBP, a RodA like protein, a PP2C type phosphatase and two FHA domain proteins. A few base pairs upstream lies a cluster of genes very similar to the cluster found in the firmicutes, coding for PriA and the other genes associated with it apart from the STPK and its associated phosphatase (figure 3.6). The apparent lack of functional linkage between the genes in the firmicute cluster can be explained via a recombination event which deleted the majority of the 'actinobacterial' genes from the 'firmicute' cluster early in the evolutionary history of this clade. Similarly, the greater distance between the actinobacterial STPK and *priA* genes can be explained by an insertion at roughly the same time as the deletion in the firmicute ancestor.

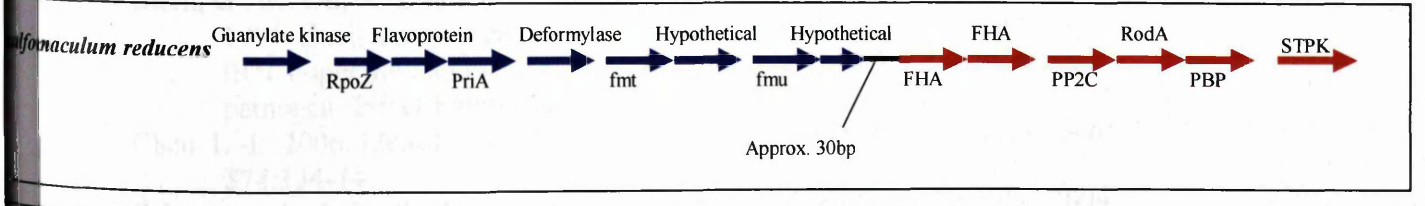


Figure 3.6: An example of *pknB* and *priA* cluster linkage from *D. reducens*. The *priA* associated genes are colured blue, *pknB* associated genes are coloured red. Gene numbers are DredDRAFT_1522 (Guanylate kinase) to DredDRAFT_1535 (STPK). STPK numbers for *D. hafniense* and *M. thermoacetica* are DSY2685 and Moth_0912 respectively.

3.4 Summary

PknB and its orthologues are present in both actinobacteria and firmicutes. These genes are almost identical to their host genomes in terms of G+C content and codon usage. Some actinobacteria carry a second *pknB* like gene. Both these genes are linked to genes coding for peptidoglycan modification enzymes. In the firmicutes the linkage between *pknB* orthologues and cell wall modification enzymes is not as clear, as the majority of sequenced genomes have a STPK and PP2C co – transcribed as an operon within a cluster of genes devoted to nucleic acid synthesis / repair. When three recently sequenced genomes are taken into account (*D. hafniense, D. reducens, M. thermoacetica*), it becomes apparent that the 'actinobacterial' *pknB* linkage is preserved in these f irmicutes, and furthermore these genomes also show linkage between the *pknB* cluster and that of *priA*.

3.5 Chapter 3 References

Blum, G., M. Ott, A. Lischewski, A. Ritter, H. Imrich, H. Tschape, and J. Hacker. 1994. Excision of large DNA regions termed pathogenicity islands from tRNA-specific loci in the chromosome of an Escherichia coli wild-type pathogen. Infect Immun **62**:606-614.

Chen, L.-L. 2006. Identification of genomic islands in six plant pathogens. Gene **374**:134-141.

Echenique, J., A. Kadioglu, S. Romao, P. W. Andrew, and M.-C. Trombe. 2004. Protein Serine/Threonine Kinase StkP Positively Controls Virulence and Competence in Streptococcus pneumoniae. Infect. Immun. **72**:2434-2437.

Han, G., and C. C. Zhang. 2001. On the origin of Ser/Thr kinases in a prokaryote. FEMS Microbiol Lett **200**:79-84.

Koonin, E. V., K. S. Makarova, and L. Aravind. 2001. Horizontal gene transfer in prokaryotes: quantification and classification. Annu Rev Microbiol **55**:709-742.

Lathe, W. C., 3rd, B. Snel, and P. Bork. 2000. Gene context conservation of a higher order than operons. Trends Biochem Sci **25**:474-479.

Leonard, C. J., L. Aravind, and E. V. Koonin. 1998. Novel families of putative protein kinases in bacteria and archaea: evolution of the "eukaryotic" protein kinase superfamily. Genome Res **8**:1038-1047.

Madec, E., A. Laszkiewicz, A. Iwanicki, M. Obuchowski, and S. Seror. 2002. Characterization of a membrane-linked Ser/Thr protein kinase in Bacillus subtilis, implicated in developmental processes. Mol Microbiol **46**:571-586.

Madec, E., A. Stensballe, S. Kjellstrom, L. Cladiere, M. Obuchowski, O. N. Jensen, and S. J. Seror. 2003. Mass Spectrometry and Site-directed Mutagenesis Identify Several Autophosphorylated Residues Required for the Activity of PrkC, a Ser/Thr Kinase from Bacillus subtilis. Journal of Molecular Biology **330**:459-472.

Munoz-Dorado, J., S. Inouye, and M. Inouye. 1991. A gene encoding a protein serine/threonine kinase is required for normal development of M. xanthus, a gram-negative bacterium. Cell **67**:995-1006.

Ogawara, H., N. Aoyagi, M. Watanabe, and H. Urabe. 1999. Sequences and evolutionary analyses of eukaryotic-type protein kinases from Streptomyces coelicolor A3(2). Microbiology **145**:3343-3352.

Petrickova, K., and M. Petricek. 2003. Eukaryotic-type protein kinases in Streptomyces coelicolor: variations on a common theme. Microbiology **149**:1609-1621.

Wolf, Y. I., I. B. Rogozin, A. S. Kondrashov, and E. V. Koonin. 2001. Genome alignment, evolution of prokaryotic genome organization, and prediction of gene function using genomic context. Genome Res **11**:356-372.

Yeats, C., R. D. Finn, and A. Bateman. 2002. The PASTA domain: a beta-lactam-binding domain. Trends Biochem Sci **27**:438.

4 Phylogenetic analysis of PknB orthologues 2: Microevolution of domains.

4.1 Structure of pknB orthologues

4.1.1 Sequence selection

As mentioned in chapter 3, the amino acid sequence of the PASTA domains from *S. coelicolor* PknB was used to perform a structural homology search using CDART (http://www.ncbi.nlm.nih.gov/Structure/lexington/lexington.cgi?cmd=rps). This search revealed many kinases sharing the general structure shown in figure 4.1 spread over the firmicutes and actinobacteria. To achieve a manageable sample, 28 firmicutes and 14 actinobacteria were examined in detail (table 4.1). In addition to the kinases, many of the firmicutes contained orthologues of a penicillin binding protein, PBP2X, from *Streptococcus pneumoniae*. This protein contains two PASTA domains at its N terminus, and is involved in localization of cell septum formation (Morlot et al. 2003). The PBP2X orthologues in the representative organisms were also analysed. See table 4.1 for the genes analysed.

A

Organism	*pknB* ortholog
Mycobacterium tuberculosis H37Rv	Rv0014c
Mycobacterium leprae TN	ML0016
Mycobacterium avium subsp. *paratuberculosis*	MAP0016c
Mycobacterium bovis AF2122/97	Mb0016c
Mycobacterium smegmatis MC2 155	MSMEGPKNB
Streptomyces avermitilis MA-4680	SAV4338
Streptomyces coelicolor A3(2)	SCO3848
Corynebacterium efficiens YS-314	CE0033
Corynebacterium diphtheriae NCTC 13129	DIP0053
Corynebacterium glutamicum ATCC 13032	CG0057
Bifidobacterium longum NCC2705	BL0589
Nocardia farcinica IFM 10152	NFA800
Leifsonia xyli subsp. xyli	LXX00210
Propionibacterium acnes KPA171202	PPA0184

B

Organism	STPK gene	*pbp2x* ortholog
Bacillus clausii KSM-K16	ABC2315	ABC2362
Bacillus halodurans C-125	BH2504	BH2573
Bacillus licheniformis ATCC 14580	BL02302	-
Bacillus anthracis str. Ames	BA4000	BA4055
Bacillus cereus ATCC 14579	BC3860	BC2662, BC3916
Bacillus thuringiensis serovar konkukian str 97-27	BT9727_3603	BT9727_2431, BT9727_3568
Bacillus subtilis subsp. subtilis str. 168	BSU15770	BSU15160

C

Organism	*pbp2x* ortholog
Clostridium acetobutylicum ATCC 824	CAC2130
Clostridium perfringens str. 13	CPE0564, CPE1863, CPE1880
Clostridium tetani E88	CTC01663
Enterococcus faecalis V583	EF0991
Lactobacillus johnsonii NCC 533	LJ0969
Lactobacillus plantarum WCFS1	LP_2200
Lactococcus lactis subsp. lactis Il1403	L89079
Listeria innocua Clip11262	LIN2145
Listeria monocytogenes EGD-e	LMO3039
Oceanobacillus iheyensis HTE831	OB1464
Staphylococcus aureus subsp. aureus MRSA252	SAR1157
Streptococcus agalactiae 2603V/R	SAG0287
Streptococcus mutans UA159	SMU.455
Streptococcus pneumoniae R6	SPR0304
Streptococcus pyogenes SSI-1	SPS0461
Symbiobacterium thermophilum	STH1205, STH2392
Thermoanaerobacter tengcongensis MB4	TTE1651

Table 4.1: Organisms and genes used in this study. A = Actinobacterial *pknB* orthologues used. In addition to the genes listed, two *Streptomyces* dcw cluster linked orthologues were used: SCO2110 and SAV6092. B = *Bacilli pknB* orthologues and *pbp2X* orthologues used. C = other firmicute *pbp2X* orthologues used.

4.1.2 Kinase domains

All PknB orthologues share the configuration of STPK at the N terminus located inside the cell, a transmembrane domain, and a number of PASTA domains (ranging from 2 – 6) located outside in the cell wall. A schematic of PknB is shown in figure 4.1. Each kinase domain shares the structure of PknB from *M. tuberculosis*, and this structure is conserved across both prokaryotes and eukaryotes (Young et al. 2003). This is confirmed using GenTHREADER (Jones 1999; McGuffin and Jones 2003) for a representative sample of PASTA domain STPKs. The kinase domains consist of two lobes, with the active site sandwiched between them. This structure is shown in figure 4.2 which also shows an ATP molecule and two Mg^{2+} ions in the cleft.

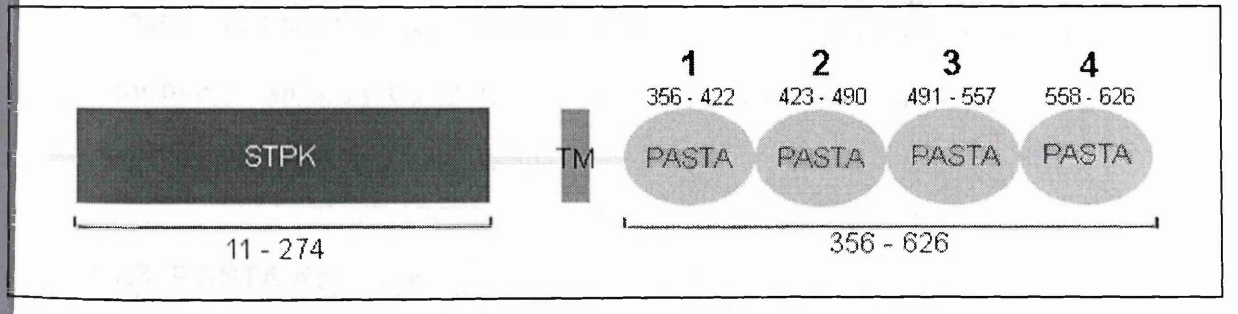


Figure 4.1: Schematic diagram of *M. tuberculosis* PknB. Numbers in bold refer to PASTA domain position, other numbers refer to amino acid number (N – C) and these indicate the domains used in phylogenetic analyses.

Figure 4.2: Representation of *M. tuberculosis* PknB kinase domain tertiary structure drawn using Cn3D 4.1. N – terminal is lowermost. Also shown are Mg^{2+} cofactors and an ATP molecule.

4.1.3 PASTA domains

The PASTA domains consist of an α – helix followed by three β – strands (figure 4.3). This structure was determined for PBP2X of *Streptococcus pneumoniae* (Pares et al. 1996), and this structure serves as the template for PASTA domains. Alignments of PASTA domain amino acid sequences show poor conservation of sequence, however structural predictions show a high degree of structural conservation. There are two small differences between PBP associated PASTA domains and those associated with STPKs. This these are a region of predominantly polar amino acids between the α – helix and the

first β – strand (figure 4.4) and a small beta sheet present at the start of the majority of the STPK associated PASTA domains that are not present in the PBP2X associated domains.

PASTA domains are found as structural repeats at the C – terminal of bacterial proteins, however are they sequential repeats? In other words, are PASTA domains more related to their immediate neighbours on the same gene as one would expect if they are involved in sensing the same ligand, or do they have different relationships?

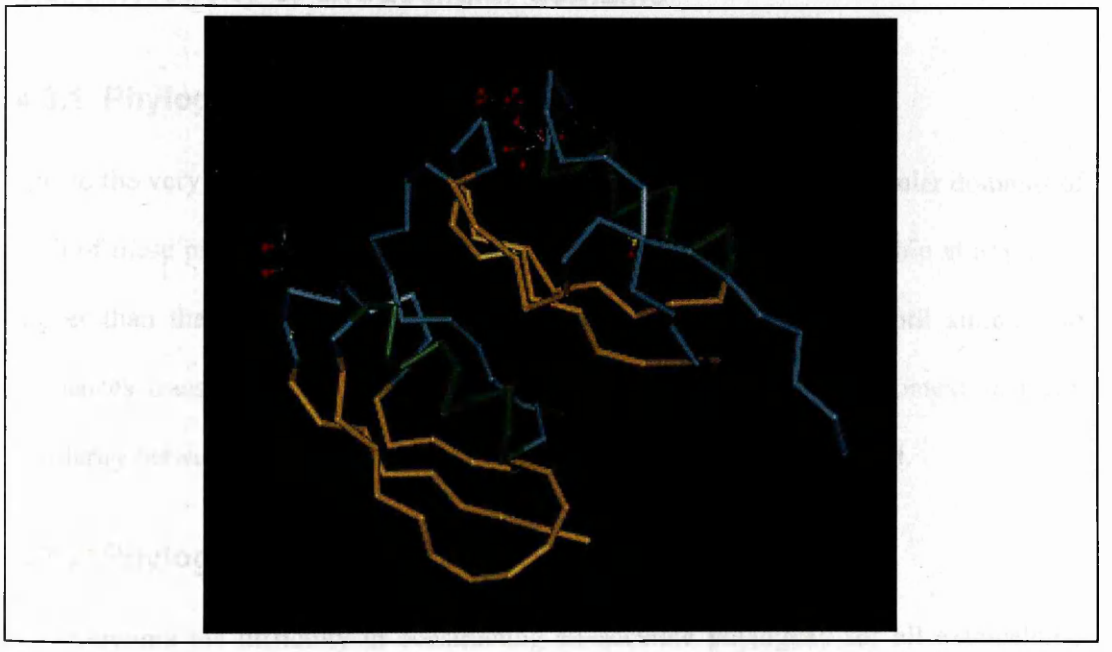
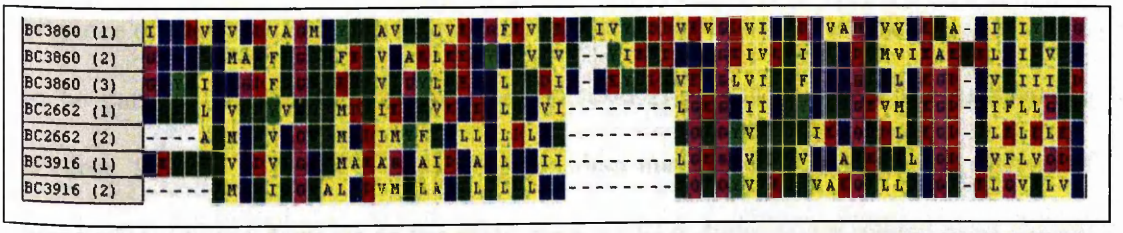


Figure 4.3: Structure of PBP2X PASTA domains drawn using Cn3D 4.1. Also shown is a molecule of cefuroxime. Helices are shown in green, beta sheets are shown in brown.



ure 4.4: Alignment of *Bacillus cereus* PASTA domains from one STPK and two PBP2X proteins. STPK is C3860, PBP2X orthologues are BC2662 and BC3916. The alignment reveals a gap of 8 mainly polar amino ids in PBP2X orthologues. This gap corresponds to a region immediately following the alpha helix.

4.2 Phylogeny of intracellular domains

In figure 3.3, the phylogeny of *pknB* orthologues was compared to that of the molecular clock genes coding for 16S rRNA and GyrA. If only the intracellular part of the genes is analysed, the phylogeny observed follows that pattern. Codon usage is the same as the genome.

4.3 Phylogeny of extracellular domains

4.3.1 Phylogeny of whole extracellular domains

Due to the very high degree of sequence divergence between the extracellular domains of each of these proteins, it is impossible to reveal a phylogenetic relationship at any level higher than that of genus with any degree of confidence. It is not until amino acid sequences translated from DNA sequences are put into a structural context that any similarity between PASTA domains from different genera can be observed.

4.3.2 Phylogeny of individual PASTA domains

To overcome the difficulty in constructing an accurate phylogeny for all extracellular domains and to investigate the relationships between PASTA domains, the individual domains were analysed. The presence of a duplicated gene in the actinobacteria and the presence of many PBP2X orthologues in firmicutes allows a more sophisticated analysis to be carried out. The results for *Mycobacteria* (figure 4.5) reveal the relationship between PASTA domains is more complex than may have been supposed. Instead of being related to the domains on the same gene, PASTA domains are related most closely to domains in an orthologue in a different organism. When organisms with a duplicated

gene are analysed, e.g. the *Streptomyces* (figure 4.6), it becomes apparent that the relationship is the same, i.e. each domain is more related to its orthologue than to its homologue. In the firmicutes the relationship shows the same positional clustering, and when the domains from *Bacilli* PASTA domains (including those from PBP2X orthologues) are examined, positional clustering of domains is seen from both kinase associated and PBP2X domains (figure 4.7).

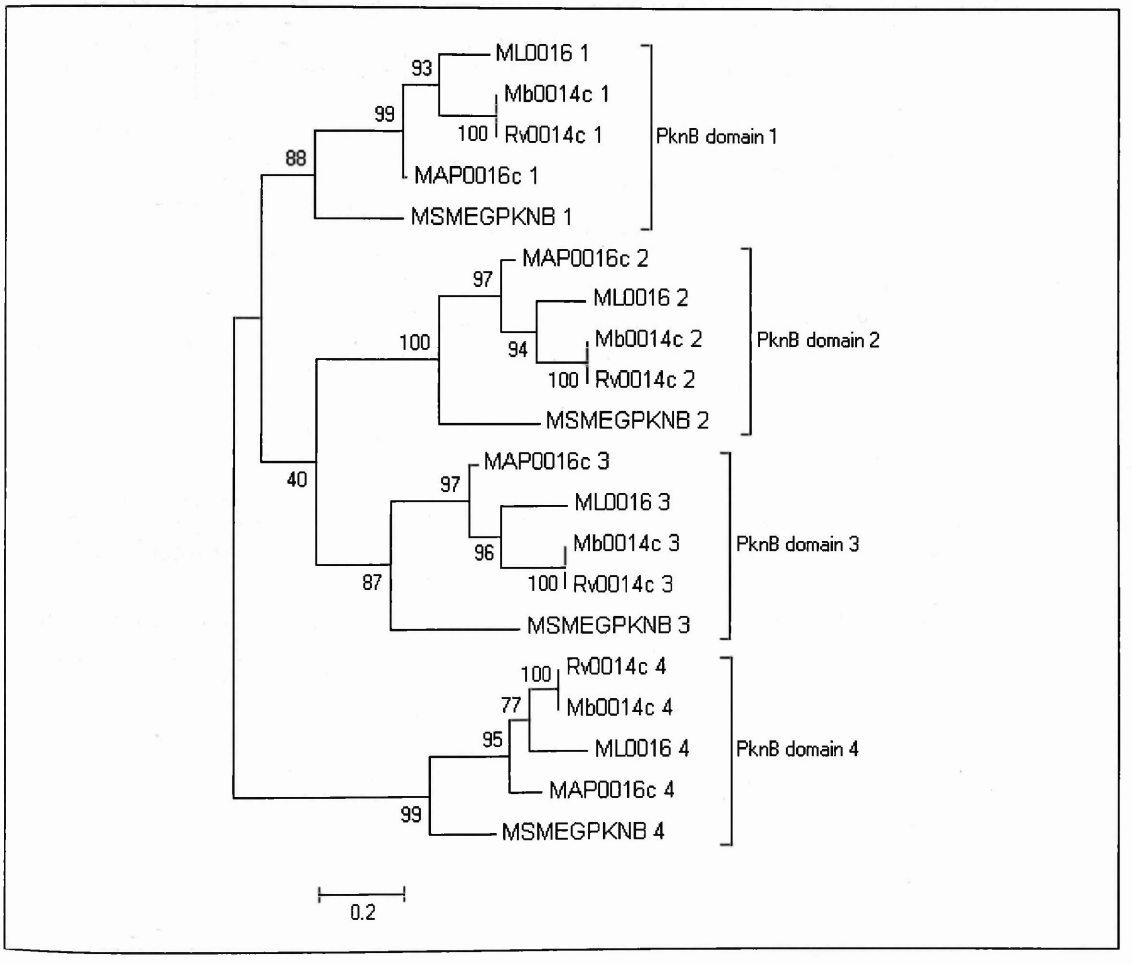


Figure 4.5: Maximum likelihood tree drawn for STPK associated PASTA domains from *mycobacteria*. Bootstrap values (percentages) from 500 replicates are shown at nodes. Numbering as per figure 4.1.

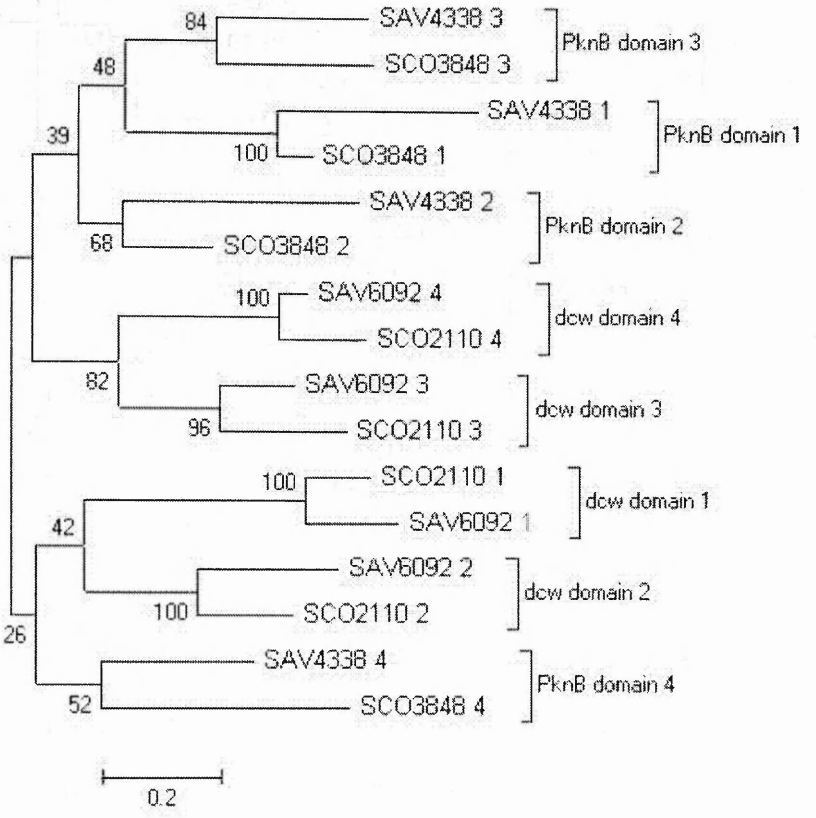


Figure 4.6: Maximum likelihood tree drawn for PASTA domains from *Streptomyces* STPKs. Bootstrap values (percentages) from 500 replicates are shown at nodes. Numbering as per figure 4.1.

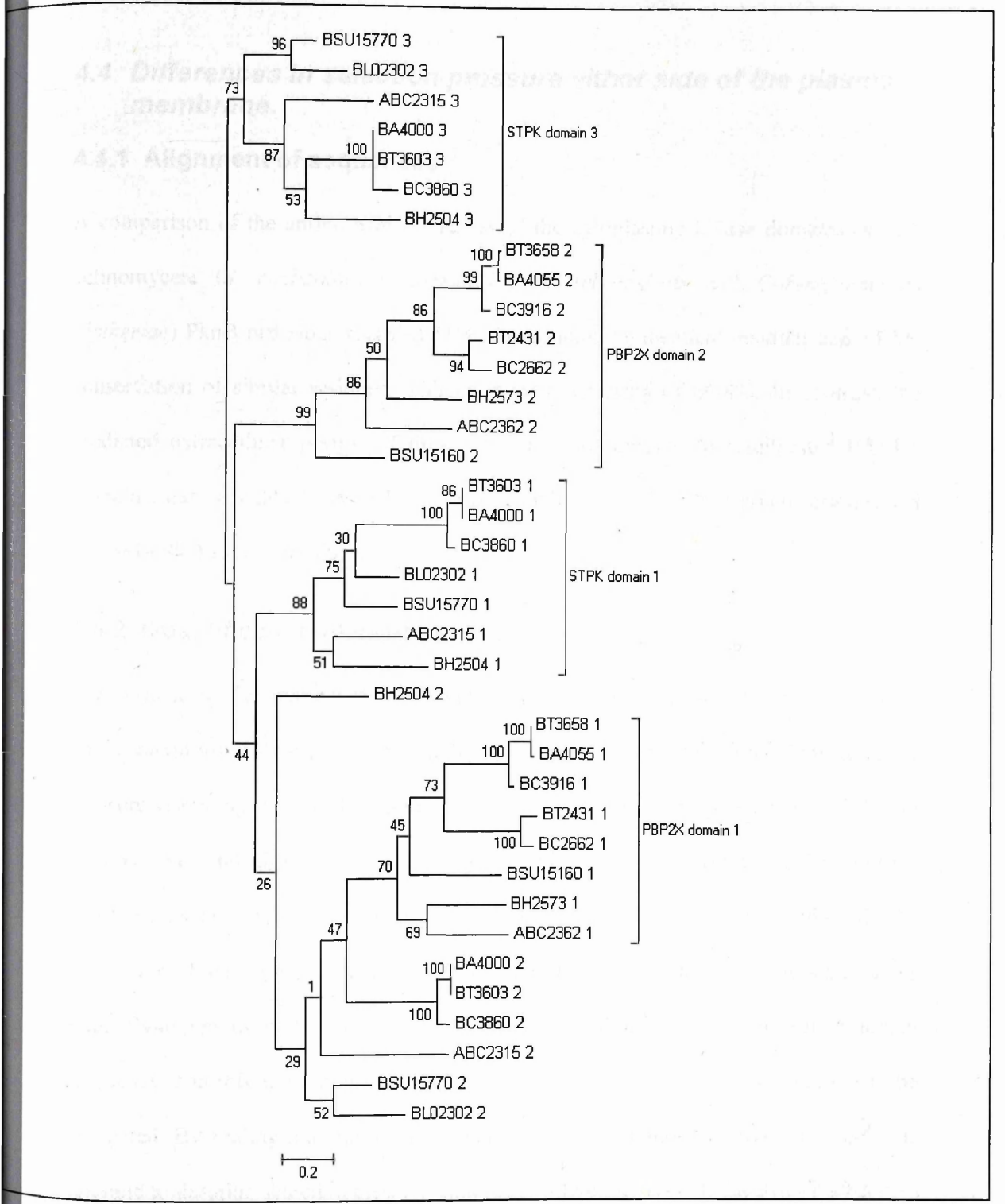


Figure 4.7: Maximum likelihood tree drawn for STPK and PBP2X PASTA domains from seven *Bacilli* species. Bootstrap values (percentages) from 500 replicates are shown at nodes.

4.4 *Differences in selection pressure either side of the plasma membrane.*

4.4.1 Alignment of sequences

A comparison of the amino acid sequences of the cytoplasmic kinase domains of four actinomycete (*S. coelicolor, S. avermitilis, M. tuberculosis* and *Corynebacterium diptheriae*) PknB orthologs revealed 55% conservation of identical residues and 11.8% conservation of similar residues, with an overall similarity of 66.8%. In contrast, the predicted extracellular portion of these proteins, consisting of four reiterated PASTA domains, has only 24.5% overall similarity, with 10.7% (1.5%) conserved residues and 13.8% (4.9%) similar residues.

4.4.2 Quantification of selective pressure

By comparing the number of synonymous DNA codon changes to the number of nonsynonymous changes, it is possible to gain a quantitative assessment of the selective pressure operating on a DNA sequence. The method used here is the Nei – Gojobori method (Nei and Gojobori 1986), using the Jukes – Cantor correction for multiple substitutions at the same site (see materials and methods). This method finds the proportion of synonymous changes (Ks) and the proportion of nonsynonymous changes (Ka). Synonymous mutation rates are taken to be indicative of the overall mutation frequency, and this assumption allows two sequences with different mutation rates to be compared. By finding the ratio of nonsynonymous : synonymous changes (Ka/Ks), it is possible to describe selective pressure either as positive (Ka/Ks > 1), neutral (Ka/Ks = 1) or purifying (Ka/Ks < 1). It is also possible to compare Ka/Ks values and make judgments on selective pressure relative to each other (Zhang and Li 2004). All Ka/Ks

values were calculated from pairwise comparisons between individual sequences and all other members of the dataset. For example, if the Ka/Ks value for SCO3848 was required, the value obtained would be the average of 13 pairwise comparisons for a data set of 14 sequences.

Sliding window analysis of the 14 *pknB* orthologues examined in this study reveals a marked difference in selective pressure either side of the membrane. The window size chosen was 72 nucleotides, with a step size of 21. This analysis is shown in figure 4.8. Furthermore, if individual PASTA domains from the same protein are compared (table 4.2) the results show that there is a low selective constraint on these domains, indeed in some cases (such as in *S. coelicolor*) there is positive selection between PASTA domains. The average score for intra gene PASTA domain comparisons is 1.02.

There are also clear differences between the degree of diversity between orthologous groups of PASTA domains on STPKs and PBP2X genes. As table 4.3 shows, PBP2X PASTA domains are significantly less diverse than the domains found in PknB – like genes, with the exception of PBP2X domain 1 and PknB domain 2.

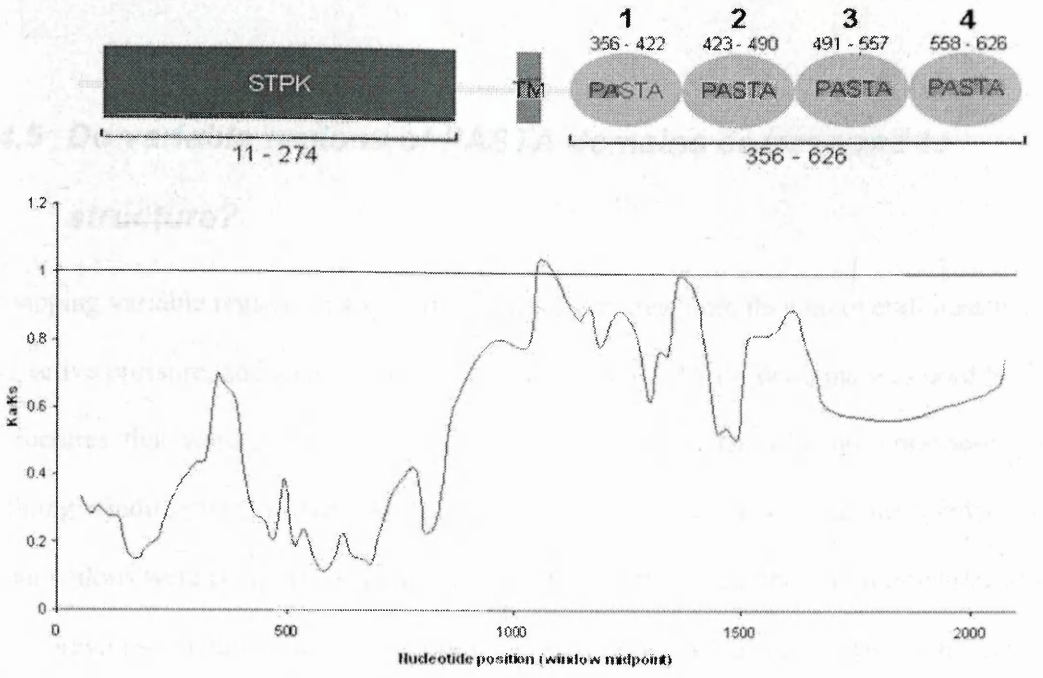


Figure 4.8: Sliding window analysis of *pknB* orthologues. Also shown is the structure from figure 3.4. Ka/Ks values are averages of pairwise comparisons between all members of the dataset.

Gene	Ka/Ks for PASTA domains
BL0589	0.81
SCO3848	1.41
SAV4338	1.00
Rv0014c	0.79
PPA0184	0.42
NFA800	1.48
ML0016	0.51
MAP0016c	1.26
Mb0014	0.79
SMEGPKNB	1.57
LXX00210	1.03
DIP0053	0.91
CG0057	0.77
CE0033	1.46

Table 4.2: Ka/Ks values for intragenetic comparisons between PASTA domains. Values shown are averages for all pairwise comparisons.

	***pbp2x* 1 (0.659)**	***pbp2x* 2 (0.524)**
***pknB* 1 (0.895)**	P<0.05	P<0.05
***pknB* 2 (0.722)**	P>0.05	P<0.05
***pknB* 3 (0.779)**	P<0.05	P<0.05
***pknB* 4 (0.859)**	P<0.05	P<0.05

Table 4.3: Comparison of average Ka/Ks values for PASTA domains at each position (figures in bold). Also shown is the result of a Mann – Whitney U test between PBP2X and STPK PASTA domains.

4.5 Do variable regions of PASTA domains correspond to structure?

Mapping variable regions in a specific sequence requires more than an overall measure of selective pressure, and so a sliding window analysis of PASTA domains was used to find structures that were mutational hotspots, or that were under selective constraint. The sliding windows used a step size of 3 nucleotides and a window of 12 nucleotides, thus four codons were compared at a time, and each window moved one codon downstream of the previous window. In each window Ka/Ks values were computed as before for pairwise comparisons between a single sequence and the remaining members of its data set. This analysis revealed positive selection primarily acting on linker regions between beta – sheets and the alpha – helix. When PBP2X domains are analysed (figure 4.9), it is apparent that these areas of positive selection correspond to areas that are mutated in penicillin resistant strains of *S. pneumoniae* (Dessen et al. 2001). In PknB orthologues the pattern of positive and purifying selection is less clear. The patterns observed for comparisons between these genes show an oscillation of positive and purifying selection as the window moves along the sequence (results not shown). These oscillations do not show a clear adherence to the structure of the PASTA domains, however the areas of positive selection generally map to linker regions between structures. Interestingly, the first beta sheet (unique to STPK associated PASTA domains) is under very tight purifying selection in the majority of comparisons. The difficulty in resolving the pattern of purifying and positive selection for STPK PASTA domains is probably due to their greater structural diversity. As has already been mentioned, the areas of positive selection appear to be constrained by tertiary structure. In STPK PASTA domains, structures do

not align as neatly as in PBP2X associated domains, as many gaps are needed to create a consensus structure. This means that as two sequences are compared, the areas of positive and purifying selection seen are unlikely to align in the same place as in a comparison between a different pair of sequences.

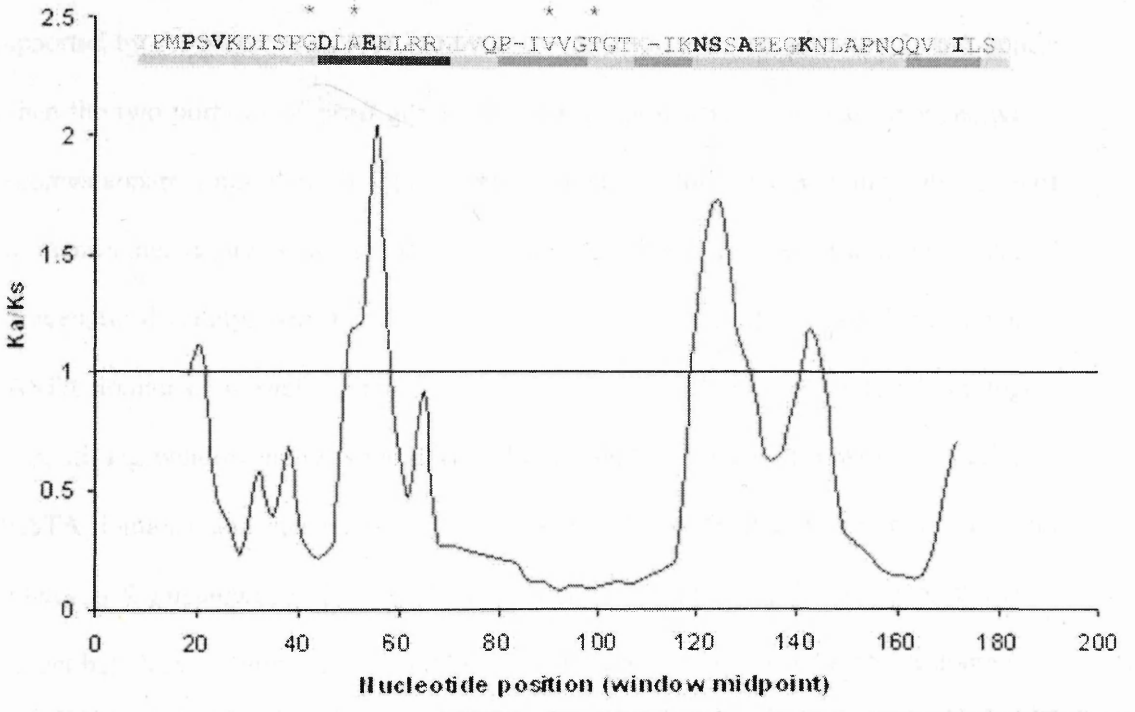


Figure 4.9: Sliding window analysis of *pbp2X* PASTA domain 1 from *S. pneumoniae*. Ka/Ks values shown are averages from pairwise comparisons between 28 sequences. Also shown is the structure and amino acid sequence of this domain. Residues in bold are mutated in penicillin resistant strains. Grayed out residues correspond to gaps in the alignment. The analysis starts at position 18 due to 'overhanging' sequences in the alignment. Residues with asterisks contact cefuroxime. Black corresponds to alpha helix, dark grey to beta sheet and light grey to linker regions.

4.6 Summary

Codon usage analysis, G+C content and phylogeny confirm that every conserved gene in the *pknB* gene cluster is of ancient origin. Furthermore, when other *pknB* orthologues are included in the analysis, they too show near identical characteristics to established molecular clocks. Genomic position and experimental data indicate that *pknB* orthologues are involved in signaling pathways involved in peptidoglycan modification, a prediction supported by the presence of PASTA domains in the extracellular portion of *pknB* genes. When the two portions of *pknB* genes are analysed from an evolutionary perspective, it becomes apparent that there is significantly different evolutionary pressure either side of the membrane. Analysis of individual PASTA domains has revealed a wide degree of intragenetic diversity, whilst intergenetic analysis has revealed positional clustering of PASTA domains, i.e. each domain is more related to an orthologue than to a homologue. Also, sliding window analysis has revealed areas under very strong positive selection in PASTA domains, and these areas map to areas which are changed in penicillin resistant isolates in *S. pneumoniae pbp2X* and these areas in turn correspond to residues shown to contact beta lactam antibiotics. A similar analysis of STPK associated PASTA domains is thwarted by the greater structural diversity between orthologues, however it is possible to say that there are areas under strong positive and purifying selection within each domain.

4.7 Chapter 4 References

Dessen, A., N. Mouz, E. Gordon, J. Hopkins, and O. Dideberg. 2001. Crystal Structure of PBP2x from a Highly Penicillin-resistant Streptococcus pneumoniae Clinical Isolate. A MOSAIC FRAMEWORK CONTAINING 83 MUTATIONS 10.1074/jbc.M107608200. J. Biol. Chem. **276**:45106-45112.

Jones, D. T. 1999. GenTHREADER: an efficient and reliable protein fold recognition method for genomic sequences. J Mol Biol **287**:797-815.

McGuffin, L. J., and D. T. Jones. 2003. Improvement of the GenTHREADER method for genomic fold recognition. Bioinformatics **19**:874-881.
Morlot, C., A. Zapun, O. Dideberg, and T. Vernet. 2003. Growth and division of Streptococcus pneumoniae: localization of the high molecular weight penicillin-binding proteins during the cell cycle. Molecular Microbiology **50**:845-855.
Nei, M., and T. Gojobori. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol **3**:418-426.
Pares, S., N. Mouz, Y. Petillot, R. Hakenbeck, and O. Dideberg. 1996. X-ray structure of Streptococcus pneumoniae PBP2x, a primary penicillin target enzyme. **3**:284-289.
Young, T. A., B. Delagoutte, J. A. Endrizzi, A. M. Falick, and T. Alber. 2003. Structure of Mycobacterium tuberculosis PknB supports a universal activation mechanism for Ser/Thr protein kinases. Nat Struct Biol **10**:168-174.
Zhang, L., and W.-H. Li. 2004. Mammalian Housekeeping Genes Evolve More Slowly than Tissue-Specific Genes
10.1093/molbev/msh010. Mol Biol Evol **21**:236-239.

5 *pknB* orthologues in *Streptomyces*: Expression and mutant phenotype

5.1 Transposon mutagenesis of SCO3848

5.1.1 Creation of a SCO3848 mutant

SCO3848 was disrupted using transposon Tn5062. Cosmid StH69 had been previously mutagenised as part of the genome – wide mutagenesis program carried out in Swansea. The insertion selected was located 372 base pairs from the ATG start codon, which ensured there would be no expression of functional SCO3848. This insertion has been named as SCH69.2.F01 on ScoDB. Following standard transposon mutagenesis methods outlined in the materials and methods, SCO3848 was disrupted *in vitro,* confirmed using restriction analysis and DNA sequencing from the ends of the transposon, before being transferred into *S. coelicolor* M145 via the *E. coli* conjugation system. Following transfer, double crossovers were selected by replica plating screening for colonies sensitive to kanamycin (selecting for the cosmid) and resistant to apramycin (selecting for the resistance gene carried on Tn5062. This method resulted in the isolation of two strains showing potential double crossovers, and the correct placement of the insertion in the *S. coelicolor* genomic DNA was confirmed by Southern blot (see figure 5.3). These strains were named 'DSCO3848(1+2)' (for Disrupted SCO3848). This analysis of insertions does not exclude the unlikely possibility of only a single crossover having occurred. The double crossover could be confirmed by using the whole of the transforming cosmid as a probe in a Southern blot, with the presence of the expected insertion pattern from a correctly inserted transposon and the absence of further cosmid – related bands confirming double crossover.

Figure 5.1: Southern blot confirmation of correct insertion of Tn5062. Lanes 1 and 5 = λ *Hin*dIII marker, lane 2 = DSCO3848(1) DNA digested *Kpn*I, lane 3 = DSCO3848(2) DNA digested *Kpn*I, lane 4 = cosmid H69.2.F01 DNA digested *Kpn*I. The probe used was DIG labelled Tn5062 and λ *Hin*dIII DNA. Arrows indicate corresponding 6.4Kb and 1.6Kb bands in both mutants and the cosmid they were made from.

The possibility of polar effects caused by this transposon insertion should not be discounted, however they do not seem to be likely for two reasons: 1) there is a strong terminator in the form of a stem – loop structure a few base pairs on from the terminator codon of SCO3848 and 2) SCO3849 is transcribed from the opposite DNA strand.

5.2 Phenotype of DSCO3848

5.2.1 DSCO3848 is affected in the timing of aerial growth and secondary metabolism.

The first impression of DSCO3848 was gleaned from its isolation on standard SFM media, and this did not indicate any immediately apparent phenotypic differences to the parental M145 strain of *S. coelicolor*. Subsequently, analysis of DSCO3848 grown on different media changed that view. The most dramatic differences were observed on solid NE media, as DSCO3848 developed aerial hyphae prematurely and produced blue pigments (likely to be actinorhodin biosynthesis related) in excess of the parental M145

strain (See figure 5.2). On media which discourages aerial growth, such as 2xYT, DSCO3848 still overproduces antibiotic when compared to the parental M145 strain (figure 5.3).

Figure 5.2. Phenotype of DSCO3848(1) bottom right, DSCO3848(2) bottom left, and the M145 parental strain top, after 48h at 30°C on NE media. The background has been left dark to emphasise aerial development.

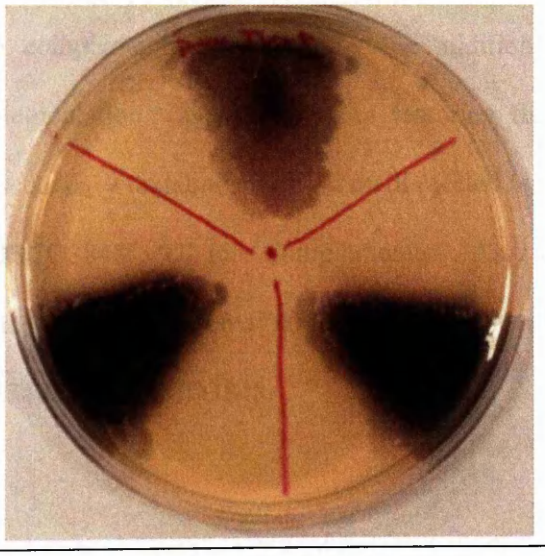


Figure 5.3, showing the increased antibiotic production of DSCO3848(1+2) on 2xYT media. The strains are arranged as per figure 6.2.

5.2.2 Function can be restored in DSCO3848 by addition of a complementing gene.

The phenotype described above was confirmed by the addition of a complementing gene contained within a pSET152 derivative, pSH152. This plasmid was made by restricting cosmid DNA containing insertion SCH69.1.E10 with *Kpn*I. This released a 5.95Kbp fragment containing SCO3848, the oppositely transcribed SCO3849 and a small portion of SCO3850. This fragment was then ligated into pAlter (also cut with *Kpn*I) to form pLK2. This plasmid was then cut with restriction enzymes *Bam*HI and *Eco*ICRI (which cuts in the *Sac*I site) and the resultant fragment was blunt – ended using SI nuclease. This fragment was then ligated into pSH152 cut with *Eco*RV. This plasmid was named pSC3848 and a diagram of its composition can be found in figure 5.4. This plasmid was introduced into both strains of DSCO3848 via the standard conjugation reaction described in the materials and methods, and colonies were selected for the disrupted copy of DSCO3848 and the complementing plasmid by the addition of apramycin and hygromycin. The complementation was successful in restoring the normal timing of aerial development and secondary metabolism on all media tested (see figure 5.5). These strains were named 'CDSCO3848', 'C' for 'Complemented'. As a control to show the effect of the vector alone, the M145 parental strain was transformed with pSH152 and plated alongside DSCO3848 and CDSCO3848.

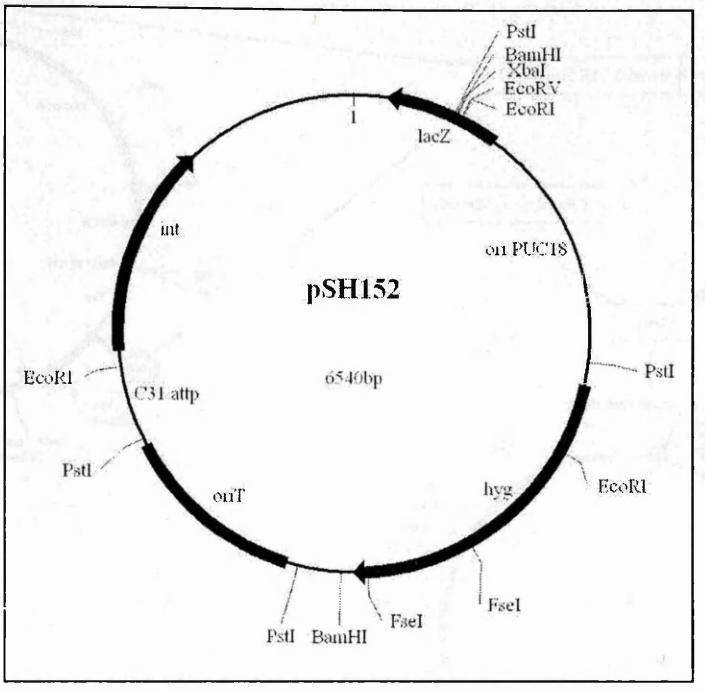


Figure 5.4a. The vector pSH152, a hygromycin resistant variant of pSET152.

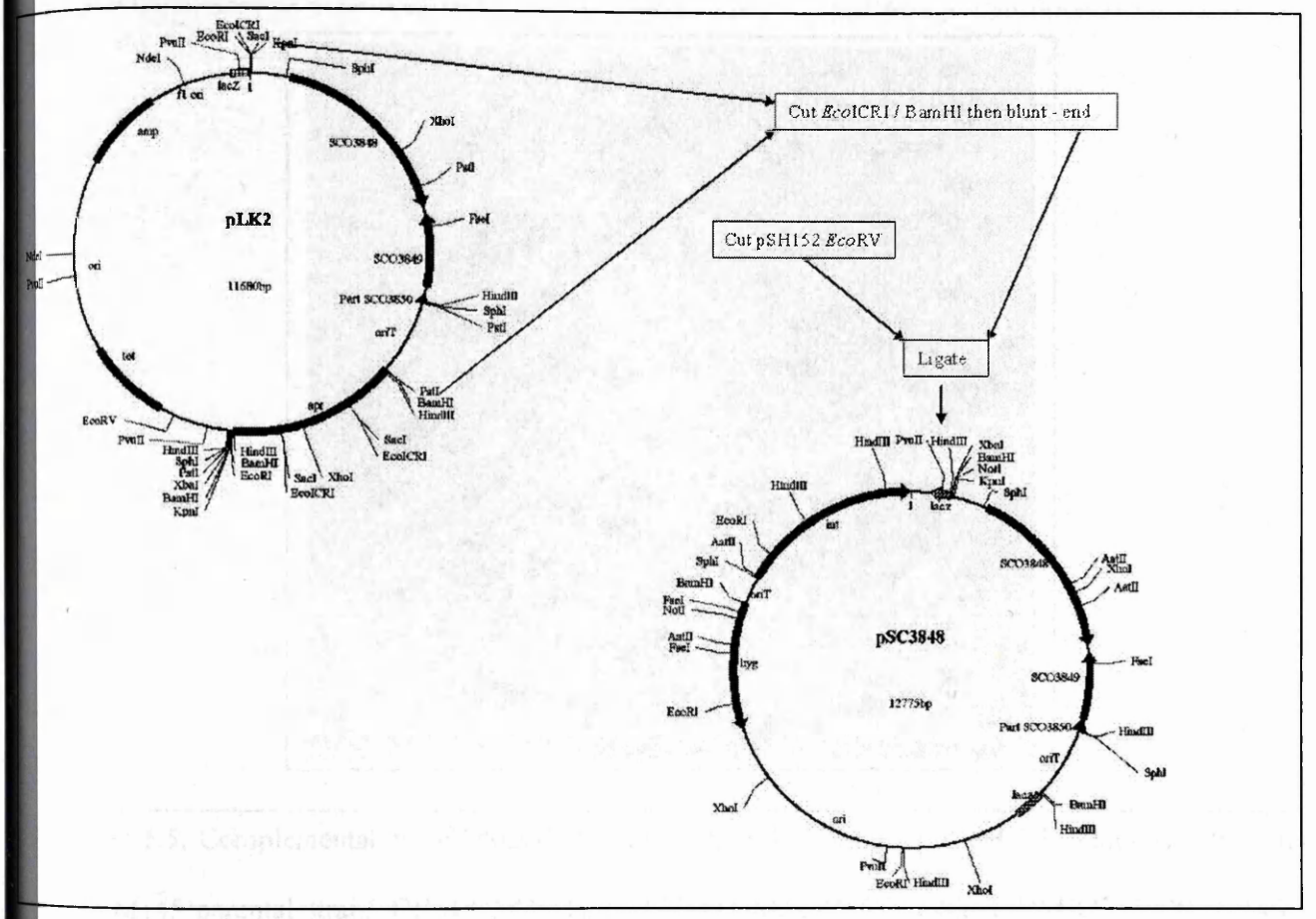
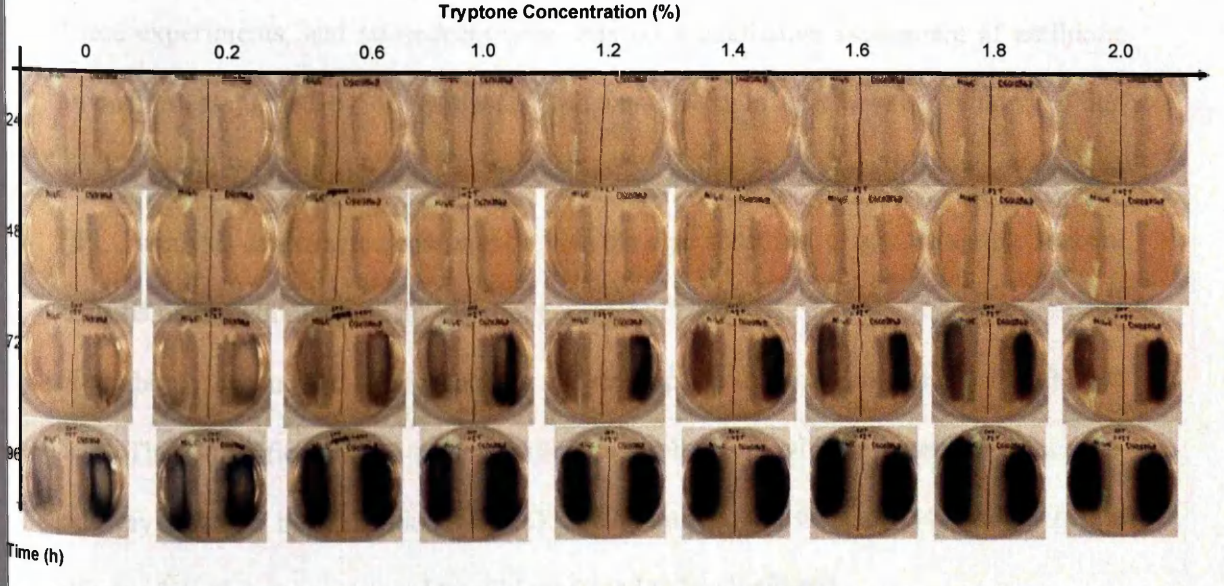


Figure 5.4b. The complementing plasmid pSC3848, created from pSH152 and a transposon insertion in SCO3850 (see text for details).

Figure 5.5. Complementation of DSCO3848 after 48h on NE buffered to pH 7.1. Clockwise from the top: M145 parental strain, CDSCO3848(1), CDSCO3848(2), DSCO3848(1), DSCO3848(2), M145 + pSH152.

5.2.3 Increasing tryptone concentration increases DSCO3848 antibiotic production.

A further insight into the function of SCO3848 can be ascertained by the examination of the effect individual media components have on the phenotype of DSCO3848. It had already been noted that 2xYT caused an increase in antibiotic production in the disrupted strain. As 2xYT contains few ingredients (tryptone, salt and yeast extract) it was decided that media containing varying concentrations of these components could be produced and the phenotypic effects examined. First to be examined was tryptone, and it was found that as the concentration rose from 0% to 2% w/v both M145 and DSCO3848 increased their antibiotic production, however DSCO3848 produced more relative to the parental strain (see figure 5.6).



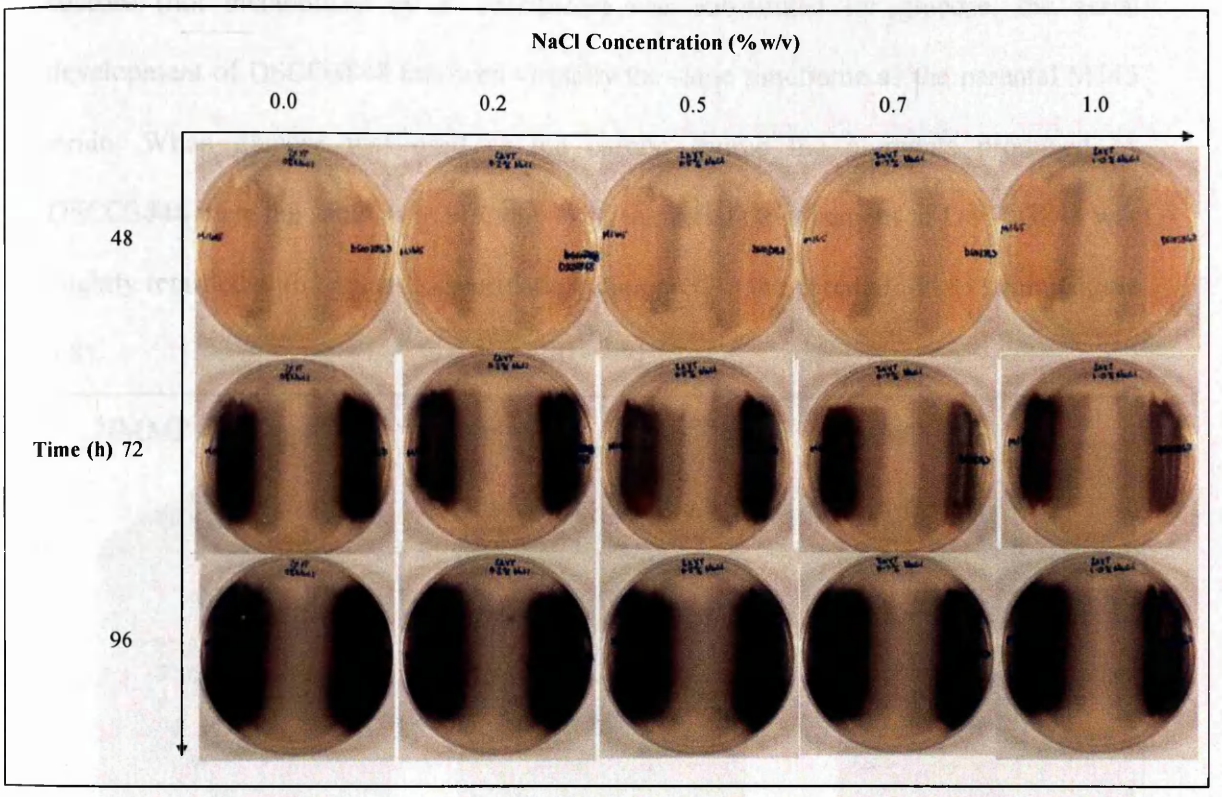
5.6. Increasing tryptone concentration is read along the x – axis. Time is read down the y – axis. Each mn of pictures for each tryptone concentration shows a time series for that particular concentration. In each re M145 is shown on the left, with DSCO3848 on the right. Basal medium is that for 2XYT.

5.2.4 Increasing NaCl concentration decreases DSCO3848 antibiotic production

The previous section noted the effect of increased antibiotic production from DSCO3848 as the concentration of tryptone increased. Commercial tryptone preparations often contain high levels of salt, so it was decided to investigate the effect increasing NaCl concentration would have on DSCO3848. Figure 5.7 clearly shows that the effect of increasing salt concentration from 0% to 1% has a negative effect on antibiotic production in DSCO3848, while the parental M145 strain remains unaffected producing the same amount of pigment at all salt concentrations.

Changing the concentration of yeast extract did not affect antibiotic production (results not shown).

These experiments, and subsequent ones, rely on a qualitative assessment of antibiotic production as judged by comparison of pigments seen in agar between wild type and mutant strains. Quantification of the antibiotics could be achieved by (for example) HPLC analysis of extracted compounds from the agar substrate, or a more simple spectral analysis of growth media. This was not considered necessary, as a qualitative assessment of antibiotic production is adequate for the description of broad phenotypic trends seen here. The quantification of excess antibiotic production will be invaluable once the pathway affected by disruption of SCO3848 is known and the enzyme kinetics of the effects of its absence compared to wild type need to be elucidated.



igure 5.7. The graph is laid out in a similar fashion to that of figure 5.6, however it is NaCl concentration hat increases from left to right on the x – axis. Time is again read from top to bottom on the y – axis. The rains in each photograph are arranged as in figure 5.6. Basal medium is 2XYT, with varying salt.

5.2.5 Addition of glucose to NMMP media delays aerial development of DSCO3848

Different carbon sources can have dramatic effects on the phenotype of *S. coelicolor*. Glucose is perhaps the most important carbon source in biology, as it is the starting molecule used in glycolysis and energy metabolism. NMMP media can be used to

investigate the effect of different carbon sources (T. Kieser 2000). It was found that when sucrose (not metabolised by *S. coelicolor*) was substituted for glucose, the aerial development of DSCO3848 followed virtually the same timeframe as the parental M145 strain. When glucose was used as the carbon source the pigments produced by DSCO3848 were the same as M145, however, the aerial development of DSCO3848 was slightly retarded with regards to sporulation compared to the parental M145 strain (figure 5.8).

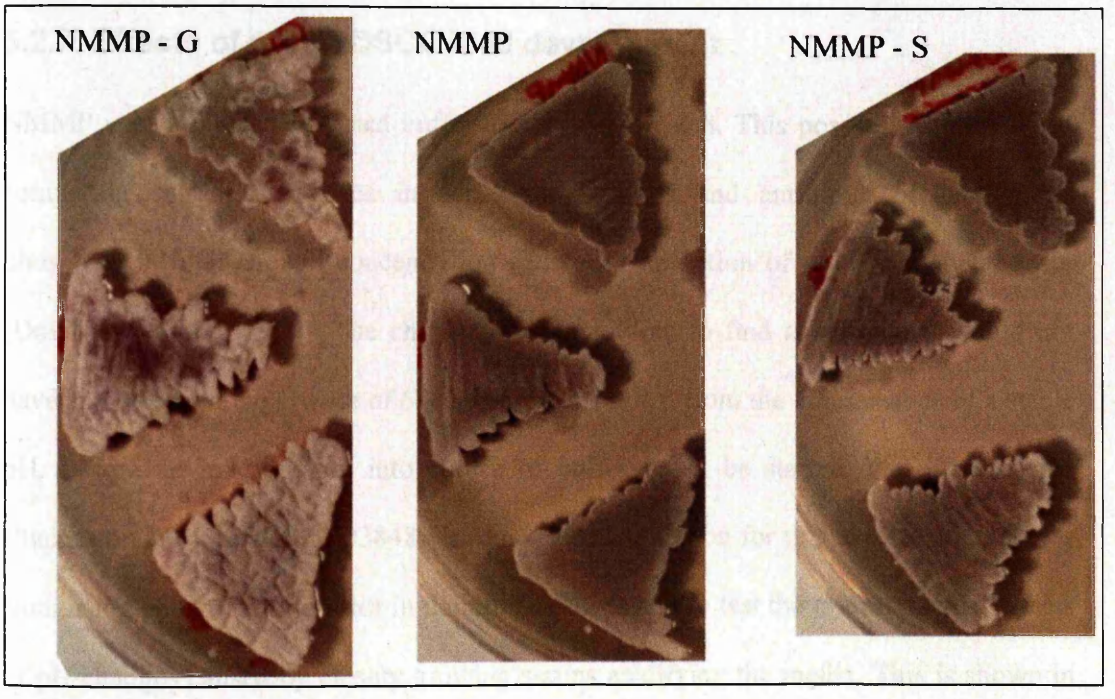


Figure 5.8: Effects of adding glucose as a carbon source to NMMP media. NMMP – G = NMMP plus 3% glucose, NMMP = NMMP with no added carbon source, NMMP – S = NMMP plus 3% sucrose as a control. In all 3 plates, from top to bottom: M145 parental strain, DSCO3848(2), DSCO3848(1). Plates were incubated for 48h at 30°C.

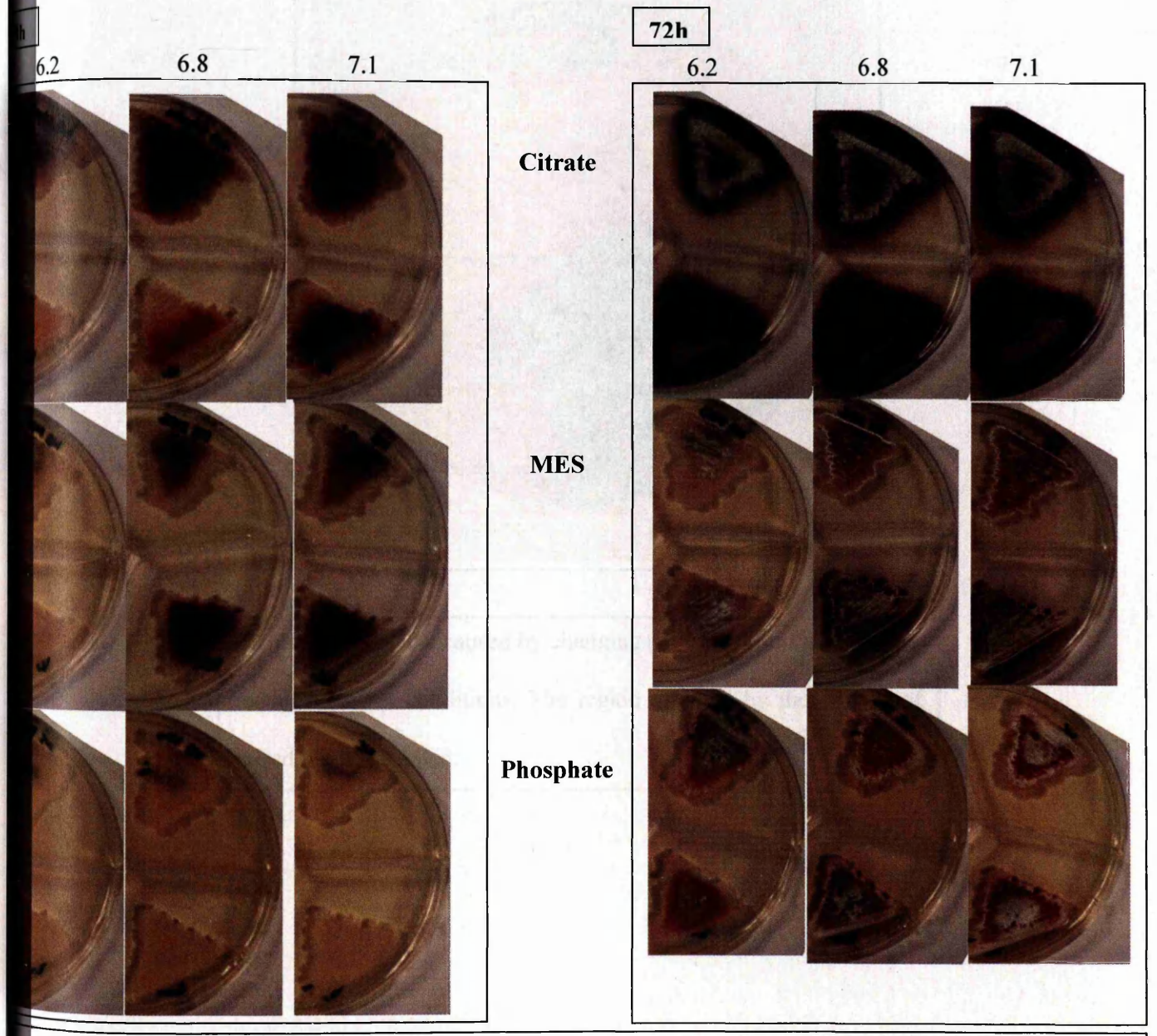
5.2.6 Effects of other carbon sources on DSCO3848 development

Other carbon sources were tried as substitutes for glucose in NMMP and minimal media (final concentrations 0.5% w/v each). These included: Mannitol; glycerol; fructose; galactose; lactose; maltose and xylose. None of the carbon sources tested caused DSCO3848 to deviate from its established phenotype of early aerial development and overproduction of antibiotic (results not shown).

5.2.7 Effects of pH on DSCO3848 development

NMMP uses a phosphate based buffer to adjust pH to 6.8. This poses a problem when comparing subtle differences in aerial development and antibiotic production, as phosphate in high enough concentration can cause inhibition of secondary metabolism (Doull and Vining 1990). The challenge was therefore to find a buffer that would not have any effect on the growth of *S. coelicolor* other than from the maintenance of a stable pH. Before an investigation into choice of buffer could be started, the effect of pH changes on M145 and DSCO3848 was analysed. The reason for this investigation comes from using phenol red indicator in unbuffered NE media to test the morphological effects of pH changes caused by closely growing strains acidifying the media. This is shown in figure 5.10. There is a 'wave' of raised pH emanating from the strain on the bottom right, and as this wave reaches the other strains it causes aerial development to start. The realisation that such dramatic pH changes were possible led to the comparison of different buffers. The results a re detailed in figure 5.9, with 10mM of each of the following buffers in SFM and NE media (composition detailed in materials and methods): Phosphate buffer as in NMMP, Citrate buffered with NaOH, MES buffered

with NaOH. The main effect shown is one of decreasing amounts of antibiotic production and aerial development as the pH decreases, however, it is the differences between buffer types that sparks most interest. As expected, the use of phosphate buffer causes inhibition of aerial development and secondary metabolism, whilst the opposite is true of the citrate buffer. MES proves to be a sensible choice of buffer when supplementing media, as it is one of Good's buffers (Norman E. Good 1966) which are considered to be relatively biologically neutral due to their reluctance to cross biological membranes.



ure 5.9: Effects of varying buffers and pH. Strains were grown on buffered NE media poured in segmented Petri dishes. and time points are indicated above the collages. In each image M145 is shown above DSCO3848.

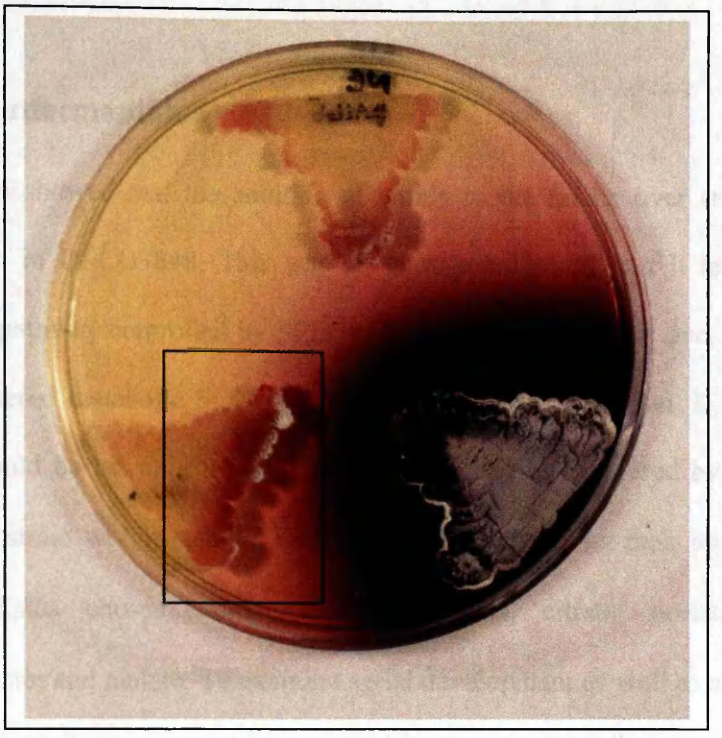


Figure 5.10: Effects caused by changing pH of media. Phenol red is red under basic conditions. The region affected by the 'wave' of raised pH is indicated.

5.2.8 Metabolic intermediates

The previous section showed that the addition of citrate to the media over stimulated antibiotic production in DSCO3848. This was an unexpected result and it led to the speculation that the pathway controlled by SCO3848 could be a metabolic one, and that its disruption had altered metabolic flux in favour of secondary metabolism. If this was the case, then it should be possible to pinpoint which pathway was altered by feeding mutant and parental strain with metabolic intermediates and observing their phenotype. Metabolic intermediates chosen were: glucose; pyruvate; citrate; acetate; α - ketoglutarate; succinate; and malate. To examine aerial development as well as antibiotic production, it was decided to use soy flour media with 10mM of the various metabolic intermediates added in the place of mannitol. The media was also buffered with MES at a final concentration of 10mM, pH adjusted to 7.1 with NaOH. With this approach it was established that there was no significant difference in either aerial development or antibiotic production between DSCO3848 and M145 for all the metabolic intermediates tested.

5.2.9 The intracellular domain only does not restore function

There has been speculation in the literature that STPKs are partially controlled by phosphorylation events at the juxtamembrane region, as well as by possible dimerisation involving the extracellular domains (Hubbard 2001; Duran et al. 2005; Wehenkel et al. 2006). To test this, a second disrupted mutant was used (created using identical methodology as for DSCO3848) . This mutant has a transposon insertion in the middle of

SCO3848 (SCH69.1.F05), and it retains the juxtamembrane and part of the transmembrane regions as well as the full kinase region. When plated on buffered NE it has the same phenotype as DSCO3848 plated on the same media (results not shown). The phenotype demonstrated by this mutant is therefore conditional on the presence of the PASTA domains.

5.2.10 Timing of transcription coincides with developmental changes.

The phenotypic changes seen thus far in DSCO3848 have been changes in the timing of aerial development at around 30 – 48h when incubated at 30°C. If the phenotype observed is a true reflection of the disruption of SCO3848, then it follows that a link between the transcription of that gene and the phenotype observed would be likely. To study the transcription of SCO3848, a plasmid (pLPK1) was constructed. This plasmid carries a fusion of the SCO3848 promoter region and the *luxA/B* genes. When the parental M145 strain is transformed with this plasmid, it allows luminescence from the *lux* system to be taken as a measure of wild type expression when compared to a control strain that has been transformed with a plasmid that contains the *luxA/B* genes but has not had the SCO3848 promoter inserted upstream. The results shown in figure 5.11 show that for all media tested, there is a peak in expression at 30h. The timing of SCO3848 expression coincides neatly with the time period during which aerial development and secondary metabolism is accelerated in DSCO3848. Interestingly, the timing of SCO3848 expression does not seem to be affected by induction of ppGpp, which points toward a different trigger of transcription (Hesketh et al. 2007).

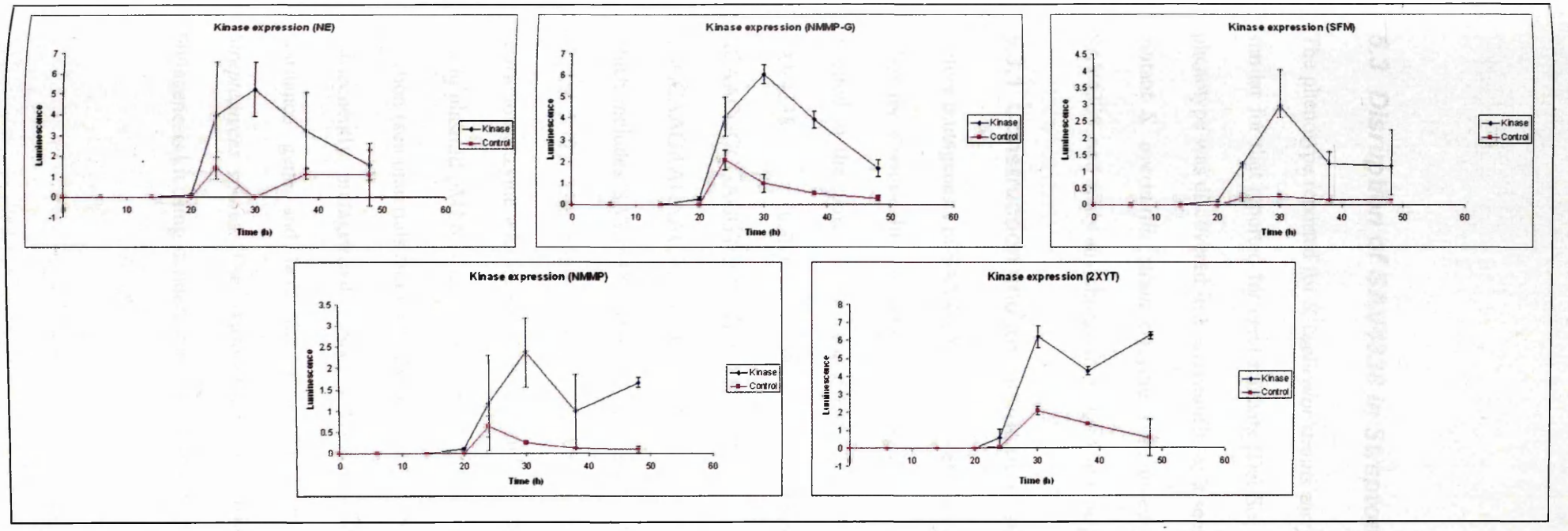


Figure 5.11: Results from analysis of kinase expression using *luxAB* system. Each graph represents a different media type indicated in the legend of each graph. Each plot contains data from control strains (representing 'basal' expression of *luxAB*). Error bars represent standard deviation.

5.3 Disruption of SAV4338 in Streptomyces avermitilis.

The phenotype reported for *S. coelicolor* strains carrying a disrupted copy of SCO3848 is similar for that reported for *crgA* mutants (Del Sol et al. 2003). The original *crgA* Whi phenotype was discovered in *S. avermitilis*, so it seemed natural to attempt to construct a mutant *S. avermitilis* strain carrying a disrupted copy of the SCO3848 orthologue, SAV4338, and study any change from the wild type phenotype.

5.3.1 Construction and mutagenesis of plasmid pMEK

Before mutagenesis of SAV4338, it was necessary to clone the gene and enough flanking DNA for a reasonable chance of obtaining a double crossover once the transposon had inserted in the gene. To do this, specific primers were designed in order to amplify SAV4338 using PCR. Forward primer sequence: ACAACGGCGAGGTGCTGGTCAGGAAGG. Reverse primer sequence: CGGCAAGAAGAAGCAGGGCGGCAATCTC. This amplifies a 3786bp fragment which includes SAV4338 and 445bp flanking DNA downstream and 1338bp upstream (figure 5.12). The primers include an *Xba*I restriction site at either end, and this restriction enzyme was used to cleave the ends of the PCR products, and to create a single cut in plasmid pME6. The PCR product and cut plasmid were used in a standard ligation reaction (see materials and methods) in order to create plasmid pMEK. This plasmid was subsequently mutagenised using transposon Tn5066, which carries a hygromycin resistance gene and *oriT* for conjugative transfer from *E. coli* to germinating *Streptomyces* spores. The transposition was carried out using the EZ:TN transposon mutagenesis kit using standard reaction conditions.

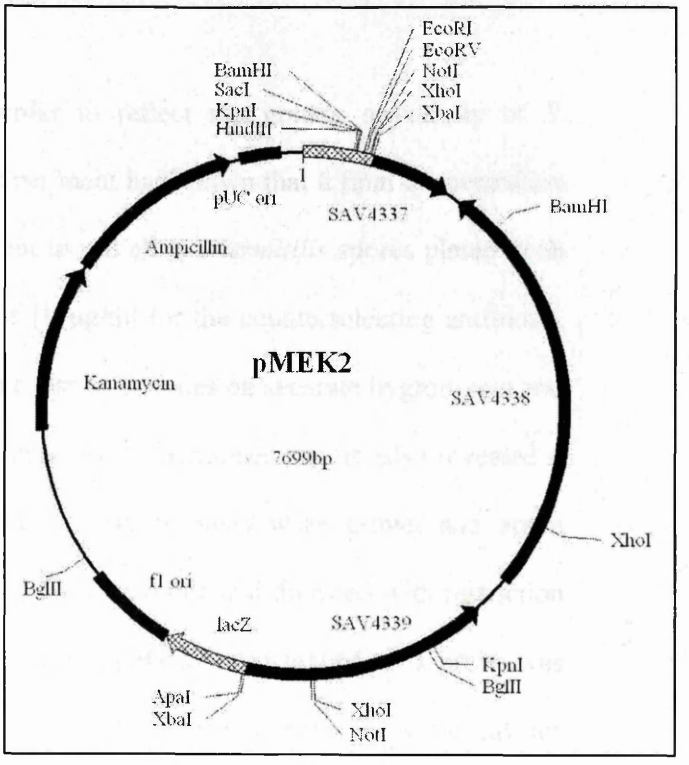


Figure 5.12: Plasmid pMEK2. The section cloned from genomic *S. avermitilis* DNA using PCR is to the right, between *Xba*I sites. The rest of the plasmid is pME6, a vector based on pCR2.1 Topo from Invitrogen.

5.3.2 Construction of DSAV4338

After mutagenesis, the mixture of mutagenised and unmutagenised plasmids in the reaction mixture was transformed into *E. coli* strain JM109 using electroporation, followed by plating on media containing hygromycin. The resistant colonies were selected and plasmids were extracted using miniprep kits. The position of the insertion was verified by restriction analysis, followed by sequencing from the end of the transposon and comparison with the published *S. avermitilis* genome sequence. Of the 10 *E. coli* colonies selected, 2 looked to have insertions in the kinase gene after restriction, and sequencing showed this was the case. The clone with an insertion closest to the start of the gene (approx 300bp from start codon) was selected for construction of the mutant strain. This plasmid was labelled pMEKtn2. pMEKtn2 was then transformed into *E. coli* ET and conjugated into *S. avermitilis* usi ng the standard conjugation method. After conjugation, the concentration of the selecting antibiotic (hygromycin) was changed from

the standard *S. coelicolor* protocol in order to reflect the greater sensitivity of *S. avermitilis* to this antibiotic. A previous experiment had shown that a final concentration of 5μg/ml hygromycin in agar was sufficient to kill all *S. avermitilis* spores plated upon it. A similar experiment showed a value of 10 μg/ml for the counterselecting antibiotic, kanamycin. Replica plating of hygromycin resistant colonies on separate hygromycin and kanamycin containing SFM plates (concentrations as mentioned previously) revealed 4 colonies that were sensitive to kanamycin. These colonies were grown and spore suspensions were made. DNA was extracted from each one and digested with restriction enzyme KpnI before Southern blot analysis using DIG labelled tn5066 as a probe was carried out in order to confirm the presence of the insertion, with the same caveats applicable as for the production of DSCO3848 (Figure 5.12).

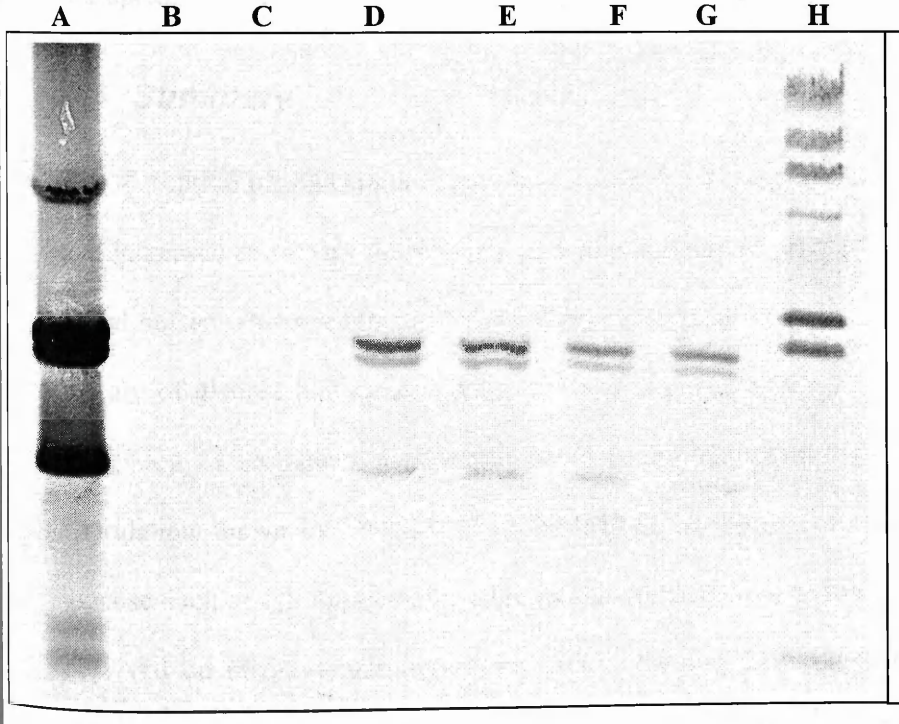


Figure 5.12: Southern blot of *S. avermitilis* SAV4338 insertions. Lane A = Plasmid DNA containing transposon in SAV4338; B = WT *S. avermitilis* DNA; C = Deliberately left blank; D – G = transposon insertions in S. avermitilis; H = λ *Hin*dIII marker.

5.3.3 Phenotype of DSAV4338

The phenotype of DSCO3848 is one of early aerial development and precocious antibiotic production on some media, whilst the onset of aerial development can be delayed in other media (eg NMMP-G). The four isolates of DSAV4338 were patched onto different culture media to see if there were any similarities or differences between two *Streptomyces* strains carrying disruptions in their *pknB* orthologues. There does not appear to be any difference between DSAV4338 and the wild type *S. avermitilis* strain on any media tested (NE, SFM, NMMP, NMMP-G, results not shown). This is surprising, considering the Whi phenotype produced when the *crgA* orthologue, SAV4331, is disrupted.

5.4 Summary

The disruption of SCO3848 results in *S. coelicolor* displaying early aerial development and increased secondary metabolism on media buffered to pH 7.1 using non – phosphate based buffer. This phenotype appears to be conditional on the carbon source used, with the use of glucose and citrate giving the most dramatic effects. The effect of phosphate repression of secondary metabolism could be the explanation for the slight delay in sporulation shown by DSCO3848 on NMMP-G, as the use of other media containing glucose such as NE supported the precocious phenotype. The same phenotype was also observed on media containing citrate, and so the link between SCO3848 and primary metabolism was investigated. No such link was found. Confirmatory evidence of the effects of disruption of SCO3848 is found when timing of its expression is found to

coincide with the time period during which DSCO3848 precociously enters its secondary metabolism and develops aerially. Disruption of the orthologous gene SAV4338 resulted in no phenotypic differences between wild type and disrupted strains.

5.5 Chapter 5 References

Del Sol, R., A. Pitman, P. Herron, and P. Dyson. 2003. The product of a developmental gene, crgA, that coordinates reproductive growth in Streptomyces belongs to a novel family of small actinomycete-specific proteins. J Bacteriol **185**:6678-6685.

Doull, J. L., and L. C. Vining. 1990. Nutritional control of actinorhodin production byStreptomyces coelicolor A3(2): suppressive effects of nitrogen and phosphate. Applied Microbiology and Biotechnology **V32**:449-454.

Duran, R., A. Villarino, M. Bellinzoni, A. Wehenkel, P. Fernandez, B. Boitel, S. T. Cole, P. M. Alzari, and C. Cervenansky. 2005. Conserved autophosphorylation pattern in activation loops and juxtamembrane regions of Mycobacterium tuberculosis Ser/Thr protein kinases. Biochemical and Biophysical Research Communications **333**:858-867.

Hesketh, A., W. J. Chen, J. Ryding, S. Chang, and M. Bibb. 2007. The global role of ppGpp synthesis in morphological differentiation and antibiotic production in Streptomyces coelicolor A3(2). Genome Biol **8**:R161.

Hubbard, S. R. 2001. Theme and Variations: Juxtamembrane Regulation of Receptor Protein Kinases. Molecular Cell **8**:481-482.

Norman E. Good, G. D. W., Wilhelmina Winter, Thomas N. Connolly, Seikichi Izawa, Raizada M. M. Singh. 1966. Hydrogen Ion Buffers for Biological Research. Biochemistry **5**:467 - 477.

T. Kieser, M. J. B., M.J. Buttner, K.F. Chater, D.A. Hopwood. 2000. Practical Streptomyces Genetics. John Innes Foundation.

Wehenkel, Fernandez, Bellinzoni, Catherinot, Barilone, Labesse, Jackson, and Alzari. 2006. The structure of PknB in complex with mitoxantrone, an ATP-competitive inhibitor, suggests a mode of protein kinase regulation in mycobacteria. FEBS Letters **580**:3018-3022.

6 SCO3843 and SCO3844: Disruption and resultant phenotype

6.1 Introduction

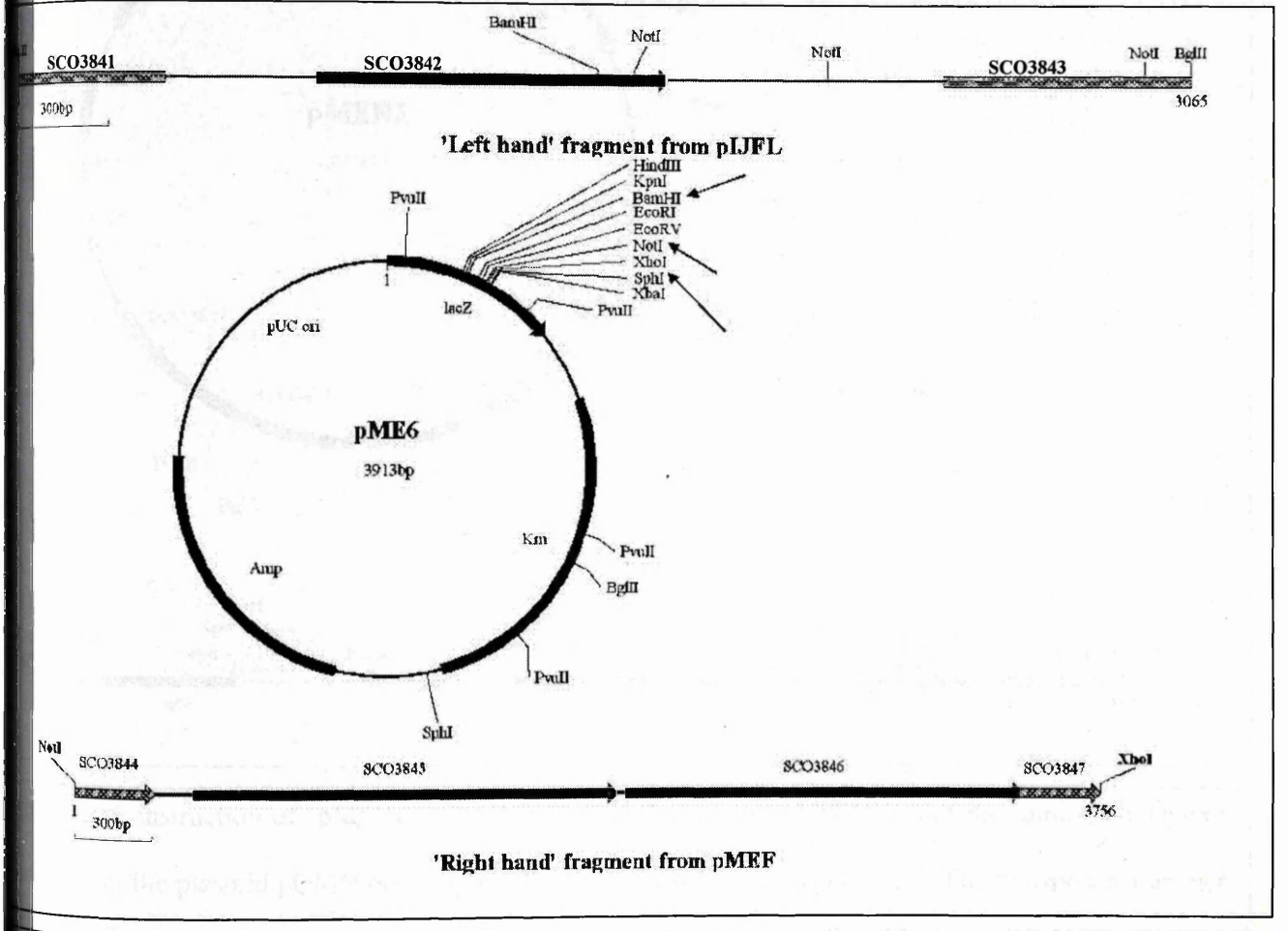
The two genes at the 'start' of the *pknB* gene cluster found in actinobacteria are members of a family of genes coding for FHA proteins recently shown to be integral to various phosphorelays in eukaryotes and prokaryotes (see introduction). The orthologue of one of these two genes has been shown to be putative acceptors of the phosphate group from the autophosphorylated PknB (Grundner, Gay, and Alber 2005). Thus far, no gene knockouts or other deletions of these genes have been reported for any of the organisms containing them. The problem of potential redundancy can be overcome by the construction of a double mutant, and so provide a more tenable position from which to predict their function in the *pknB* gene cluster.

6.2 Construction of a SCO3843 and SCO3844 double mutant

6.2.1 Plasmid construction

The creation of a double disruption of SCO3843 and SCO3844 required enough DNA either side of disrupted copies of the two genes. To do this, relevant fragments were cloned from plasmid pMEF and from cosmid H69. These fragments were inserted either side of a multiple cloning site on plasmid pME6. To allow transfer from *E. coli* to *S. coelicolor* in a standard conjugation reaction, a copy of Tn5066 was inserted between the two fragments at an *Eco*RV position. The transposon provided a convenient way of

introducing both an origin of transfer and a selectable antibiotic resistance marker (in this case, hygromycin). For details of the construction of the plasmid used in the final conjugation, pMERL66, please see figures 6.1 and 6.2. The end result was to have the FHA domains of SCO3843 and SCO3844 (as predicted by SMART (Letunic et al. 2004)) deleted and replaced with a transposon carrying hygromycin resistance and *oriT*.



igure 6.1: Construction of pMERL. The plasmid pME6 was first cut with *Xho*I and *Not*I, as was plasmid MEF. Upon purification of the relevant bands from agarose gels, the 'right hand' fragment was ligated into ME6, yielding the interim plasmid pMER. Following confirmation of successful ligation using restriction lysis, the left hand arm was cloned into position by using the fact that *Bam*HI and *Bgl*II create fragments identical sticky ends. Thus, the fragment from pIJFL could be cut out using *Bgl*II and ligated into pMER ing the *Bam*HI site on the multiple cloning site.

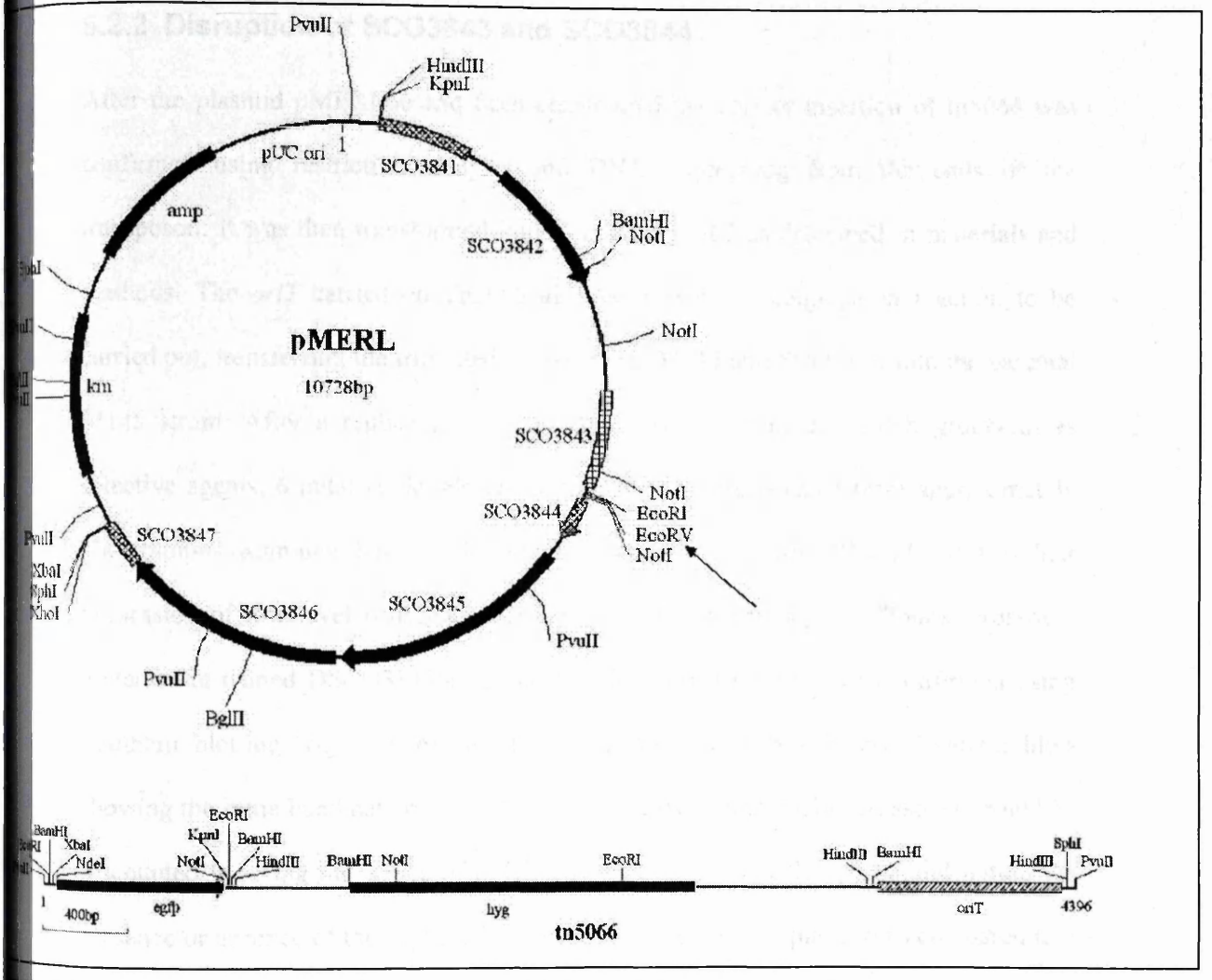
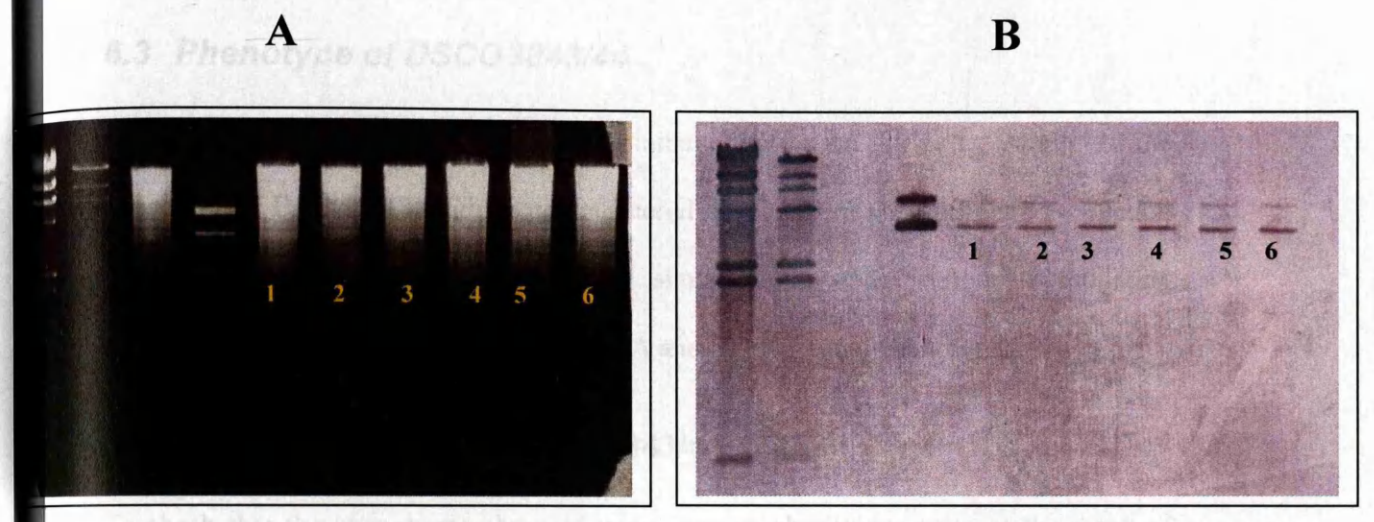


Figure 6.2: Construction of pMERL66. pMER was restricted using *Eco*RV and at the same time Tn5066 was cut from the plasmid pQM5066 using *Pvu*II and purified from an agarose gel. The transposon was then blunt – end ligated into pMERL to form pMERL66. The orientation of the transposon in pMERL66 was checked using restriction analysis and sequencing, and clones containing both orientations were isolated.

6.2.2 Disruption of SCO3843 and SCO3844

After the plasmid pMERL66 had been constructed the correct insertion of tn5066 was confirmed using restriction analysis and DNA sequencing from the ends of the transposon. It was then transformed into *E. coli* ET12567 as described in materials and methods. The *oriT* carried on Tn5066 allowed a standard conjugation reaction to be carried out, transferring the truncated copies of SCO3843 and SCO3844 into the parental M145 strain. After a replica plating procedure using kanamycin and hygromycin as selective agents, 6 putative double crossover mutants were isolated from approximately 200 colonies examined. Upon further examination, 4 of these were discarded due to their possession of low level resistance to kanamycin. The remaining two double crossover mutants are named DSCO3843/44(1 and 5). The correct insertion was confirmed using Southern blotting (fig 6.3). As in Chapter 5, the possibility of t hese Southern blots showing the same band pattern for both single crossovers and double crossovers could be discounted by using the whole of pMERL66 as a DIG – labelled probe and noting the presence or absence of the supposedly degraded section of this plasmid as compared to a positive control.



ure 6.3: Southern blot confirming correct insertion of Tn5066 in M145 chromosome. A = Agarose gel used make Southern blot (B). Lanes in both A and B from left to right: λ*Hin*dIII marker; λ*Hin*dIII marker diluted 0; M145 *Sac*I digested; pMERL *Sac*I digested; DSCO3843/44 (1 – 6) *Sac*I digested. Subsequently, clones 2, 4 and 6 were found to be single crossovers and were discarded. The probe used in the Southern blot (B) is 5066 labelled with DIG, mixed with similarly labelled λ DNA.

6.3 Phenotype of DSCO3843/44

At first sight the phenotype of the double mutant showed an intriguing parallel to the phenotype of DSCO3848, with precocious aerial development and antibiotic production on buffered NE media. After this initial similarity, however, further investigation revealed several differences between the FHA and STPK mutant phenotypes.

6.3.1 Complementation of DSCO3843/44

To check that the phenotypes observed were genuine phenotypes and not the result of unexpected interactions between the transposon and the host, a complementing plasmid was made, pSEF2, using a *Pvu*II / *Eco*RV fragment containing SCO3843 and SCO3844 from pMEF. This is a derivative of the commonly used plasmid pSET152, carrying apramycin resistance. pSEF2 was transformed into both mutants, and pSET152 was transformed into the parental M145 strain. These two strains were created to make sure that the host plasmid itself had no phenotypic effects (figures 6.4a and 6.4b). The phenotype shown by DSCO3843/44 is twofold. Firstly, there is accelerated aerial development at around 36 hours incubation at 30°C. This early aerial development proceeds to full sporulation, which is again earlier than in the M145 parental strain (figure 6.4a). The second part of the phenotype is overproduction of antibiotic, as shown in figure 6.4b. The phenotypes displayed are not affected by insertion of an empty plasmid (pSET152), however when this plasmid has a complementing fragment inserted in it to form pSEF, the parental phenotype is restored.

ure 6.4a: Complementation of DSCO3843/44. Clockwise from the top: **A**) M145, **B**) DSCO3843/44(1) + T152, C) DSCO3843/44(5) + pSET152, **D**) DSCO3843/44(1) + pSEF2, **E**) DSCO3843/44(5) + pSEF2, SCO3843/44(1), **G**) DSCO3843/44(5), **H**) M145+pSET152. The media used was SFNAG buffered to 7.1 using MES buffer. The time point shown is 36h, which emphasises the precocious nature of mutant al development.

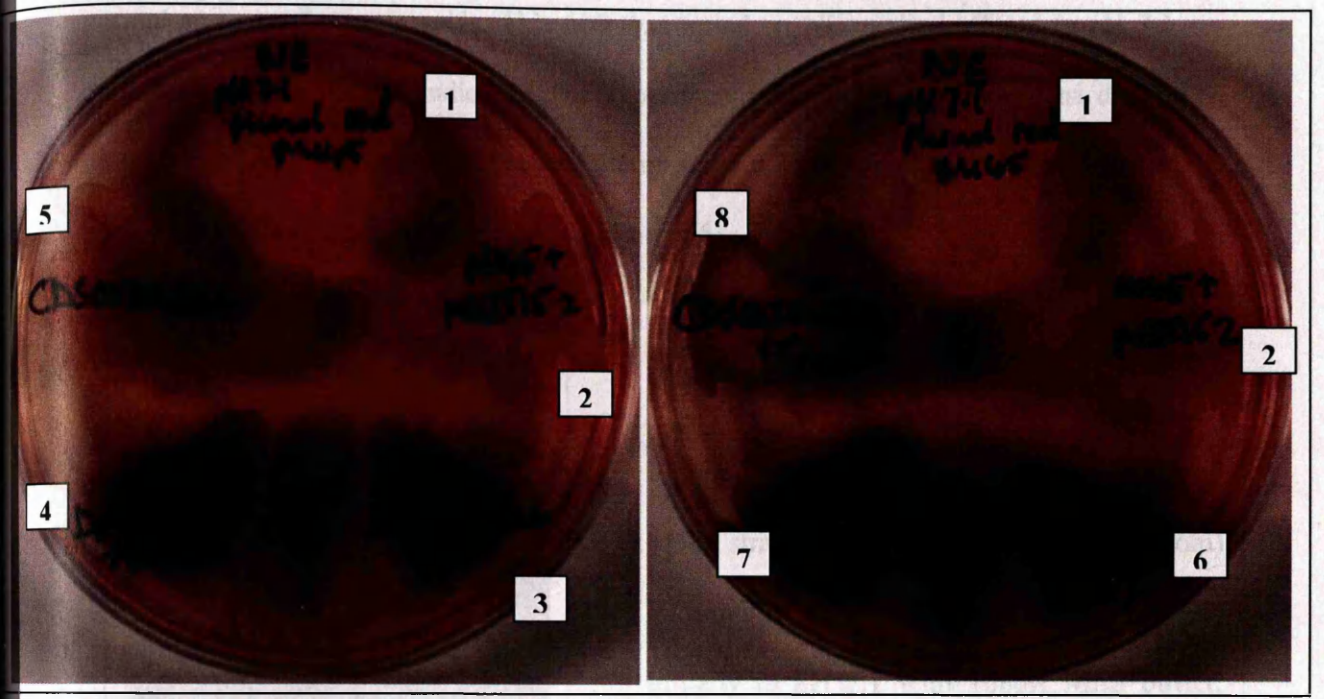


Figure 6.4b: Complementation of DSCO3843/44 regarding antibiotic production. Media is NE buffered to pH 7.1 with 10mM MES buffer and phenol red pH indicator, time point is 48h after inoculation. 1 = M145, 2 = M145 plus pSET152, 3 = DSCO3843/44(1), 4 = DSCO3843/44(1) plus pSET152, 5 = DSCO3843/44(1) plus pSEF, 6 = DSCO3843/44(5), 7 = DSCO3843/44(5) plus pSET152, 8 = DSCO3843/44(8) plus pSEF.

6.3.2 DSCO3843/44 pH effects

As has been mentioned previously in chapter 5, the effect of pH changes in the culture media can have a dramatic effect on the morphology of a *Streptomyces* patch when grown on solid media. DSCO3843/44 is not only affected by changes in pH, it does not display the same ability to change the pH of its environment as the parental strain. Phenol red indicator changes colour from yellow to red as pH increases from 6.6 to 8.0. Thus, DSCO3843/44 does not appear to have the same ability to raise the pH of its surroundings, and this may be due to altered metabolic flux producing a larger than

normal amount of organic acids. This trait was tested on NE media buffered to pH 7.1 containing phenol red pH indicator. The experiment was performed using petri dishes divided into three chambers that prevent the movement of molecules from one chamber to the other. The effect is demonstrated in figure 6.5.

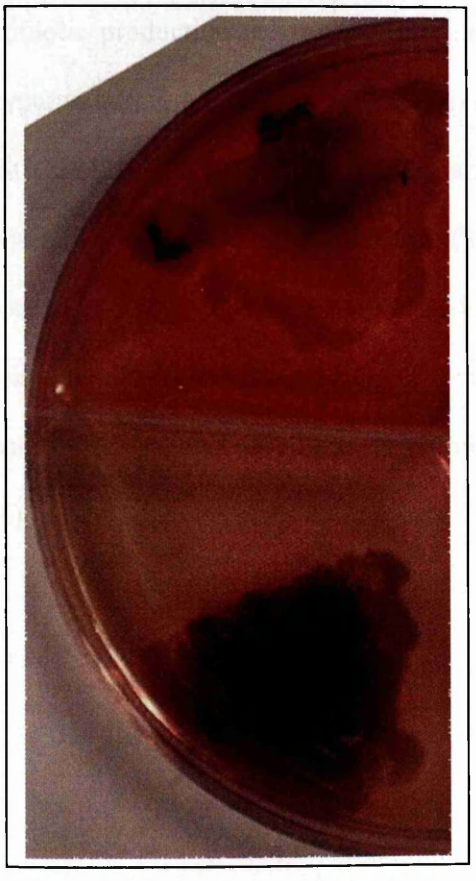


Figure 6.5: Comparison between parental M145 strain (top) and DSCO3843/44 (bottom) in terms of pH. Media used is NE buffered initially to pH 7.1 with 10mM MES buffer, also added is phenol red as a pH indicator. The plate was incubated at 30 ºC for 48h before this picture was taken.

6.3.3 The effects of phosphate and carbon source on DSCO3843/44

Unlike DSCO3848, DSCO3843/44 is not repressed in terms of antibiotic production when grown on media containing 10mM phosphate buffer at varying pH. Instead, antibiotic production is greater than that of the wild type in all media. On NE media (tryptone based, glucose as carbon source) supplemented with 10mM phosphate buffer, DSCO3843/44 shows a bald phenotype. This is rescued when it is grown on media containing mannitol as a carbon source (SFM). The effects of 10mM citrate are similar to those of 10mM phosphate, however the buffer type is important in terms of repression of aerial growth, as using one of Good's buffers (Norman E. Good 1966), MES, at 10mM concentration allows aerial growth on media containing both carbon sources (see figure 6.6)

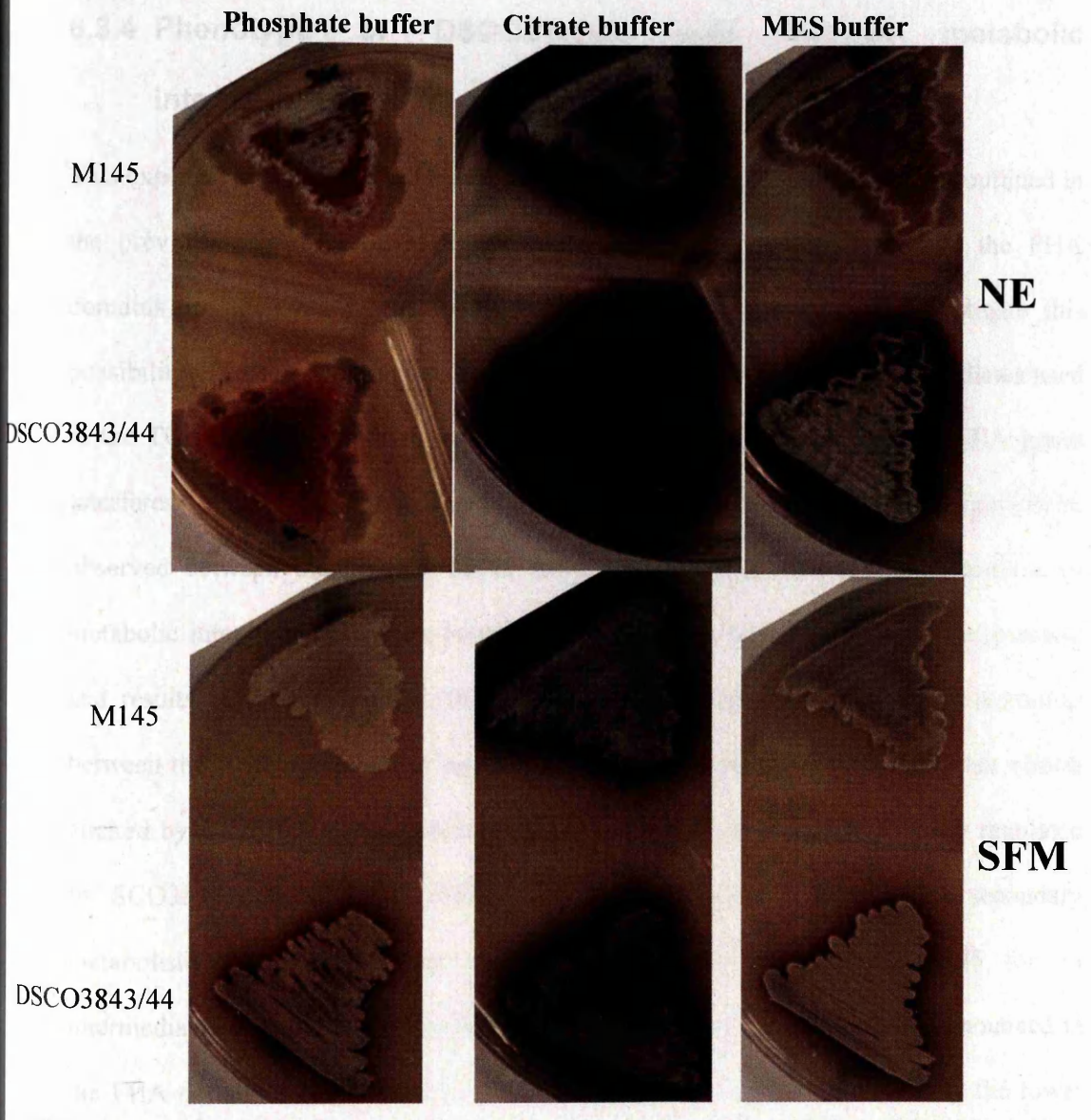


Figure 6.6: Effects of buffer and carbon source on DSCO3843/44. All images were taken after 72h incubation at 30°C. The concentration of all buffers is 10mM, for their composition, please see materials and methods.

6.3.4 Phenotype of DSCO3843/44 with various metabolic intermediates

This experiment was prompted by the phenotype observed during experiments outlined in the previous section. This prompted speculation, as for DSCO3848, that the FHA domains are involved in the regulation of primary metabolism. To investigate this possibility, 'feeding' experiments were carried out with key metabolic intermediates used in the TCA cycle. If the disruption of the signalling cascade involving the FHA genes interferes with the TCA cycle, then it seems likely that a difference in phenotype will be observed between the parental M145 strain and DSCO3843/44 as concentrations of metabolic intermediates change between strains. Figures 6.7 and 6.8 show the principle and results of this experiment. In general terms, it seems that there is a relationship between the FHA domains and regulation of primary metabolism. The different effects elicited by the different intermediates makes pinpointing a particular pathway regulated by SCO3843 and SCO3844 difficult, although the general trend is for secondary metabolism to be much more vigorous in DSCO3843/44 than in M145 for all intermediates tested. The overproduction of antibiotics is very much more pronounced in the FHA domain mutants than in DSCO3848, and this may be explained by the lower number of potential substitute genes, 7 FHA domains in the *S. coelicolor* genome compared to approximately 40 STPK genes. It is interesting to note the stimulation of antibiotic in M145 when it is grown on media supplemented with malate, citrate and succinate. The phenotypes of mutant and parent become much more similar on media supplemented with these intermediates, and this contrasts with the lack of antibiotic

production on unsupplemented media, and on media supplemented with pyruvate, acetate and α-ketoglutarate. The implication is that the FHA domains are involved in control of metabolic flux involving the intermediates on which DSCO3843/44 overproduces but the parental M145 strain does not.

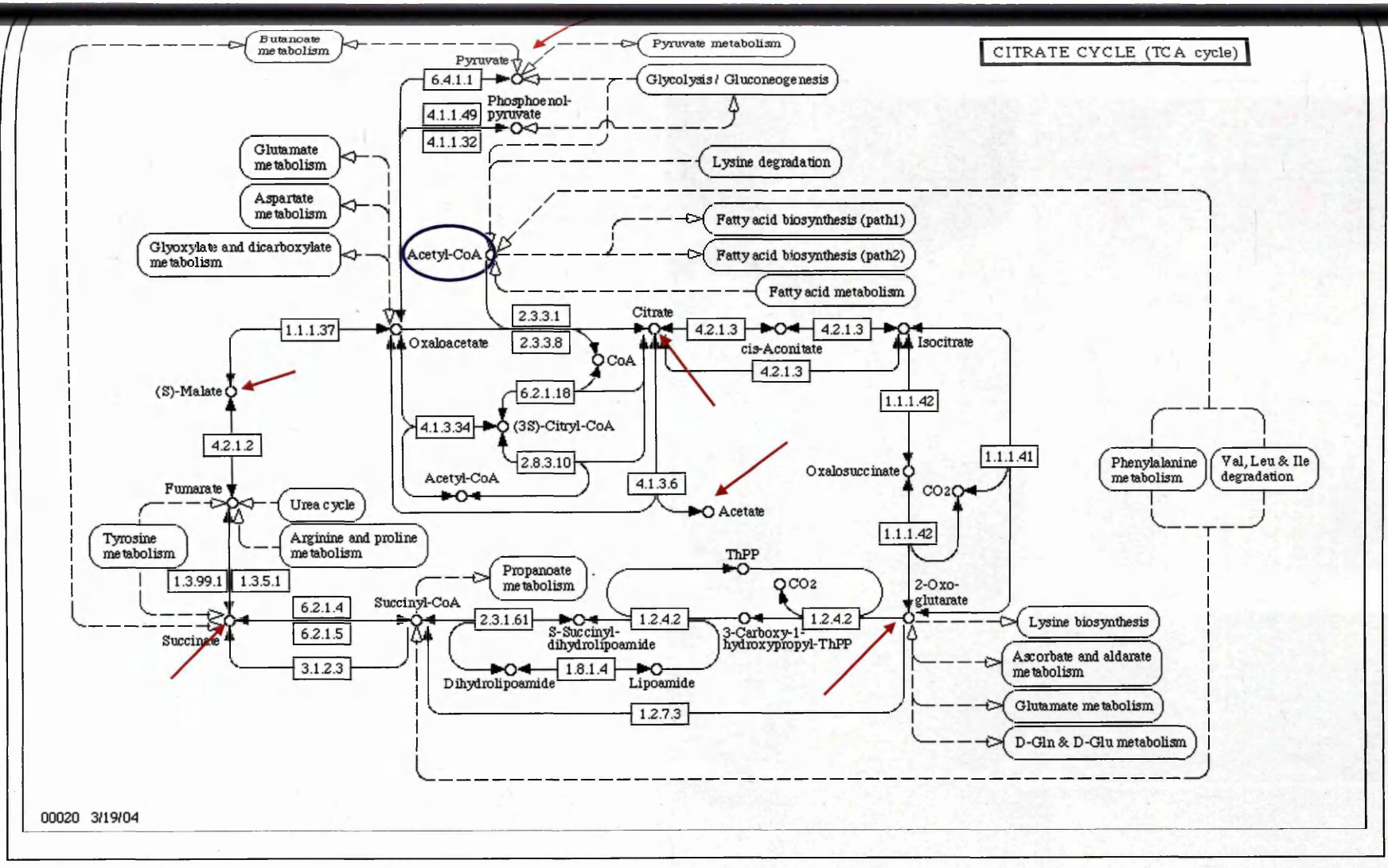


Figure 6.7: TCA cycle as displayed by KEGG (Kanehisa et al. 2006). The intermediates which were artificially increased in the feeding experiments mentioned in section 6.3.3 are indicated (red arrows). Acetyl Co-A is also highlighted as this is a precursor for polyketides.

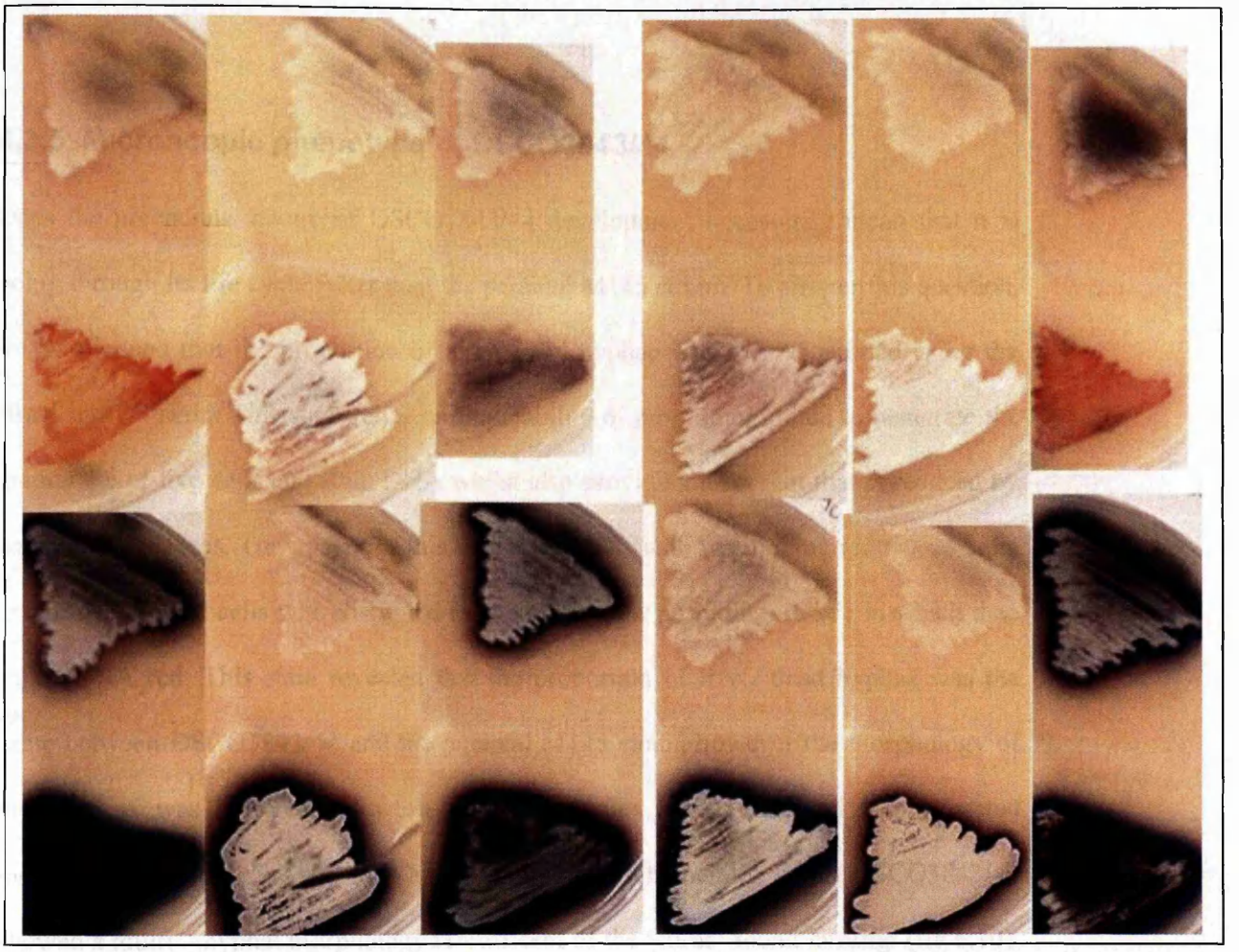


Figure 6.8: Results of feeding DSCO3843/44 with metabolic intermediates. Media used was 2% soy flour media supplemented with 10mM MES buffer (pH 7.1) and 10mM various metabolic intermediates.

6.3.5 Microscopic phenotype of DSCO3843/44

Does the precocious nature of DSCO3843/44 development necessarily mean that it is going through its life cycle faster than the parental M145 strain? To answer this question, it was planned that the proportion of live / dead hyphae would be examined using the 'BacLight™' stain. This stain relies on the inability of propidium iodide to penetrate the membrane of live cells and stain DNA whilst also providing a reagent that is reduced by bacterial reductases. The upshot of this is that live cells with functioning reductases stain green, whilst dead cells with disrupted membranes allow propidium iodide in which then stains DNA red. This stain revealed that the proportion of live / dead hyphae was the same between DSCO3843/44 and the parental M145 strain, however the morphology of the hyphae was different between the two strains. M145 showed the expected morphology of individual substrate hyphae, with normal branching, whilst DSCO3843/44 showed a multi – hyphal morphology reminiscent of the 'cords' found in fungi (fig 6.10). Regarding the live / dead state of the cells, substrate hyphae died as the colonies got older (i.e. more stained red, fewer stained green) however a surprising result was noted for the spores of M145. It was expected that most of the spores would stain green, as after all they are the route by which *S. coelicolor* is thought to favour reproduction. Instead, many of the spores in the parental strain were stained red, indicating that they had disrupted membranes and were therefore dead (figure 6.9). This was also true for DSCO3843/44, although fewer spore chains were observed.

A

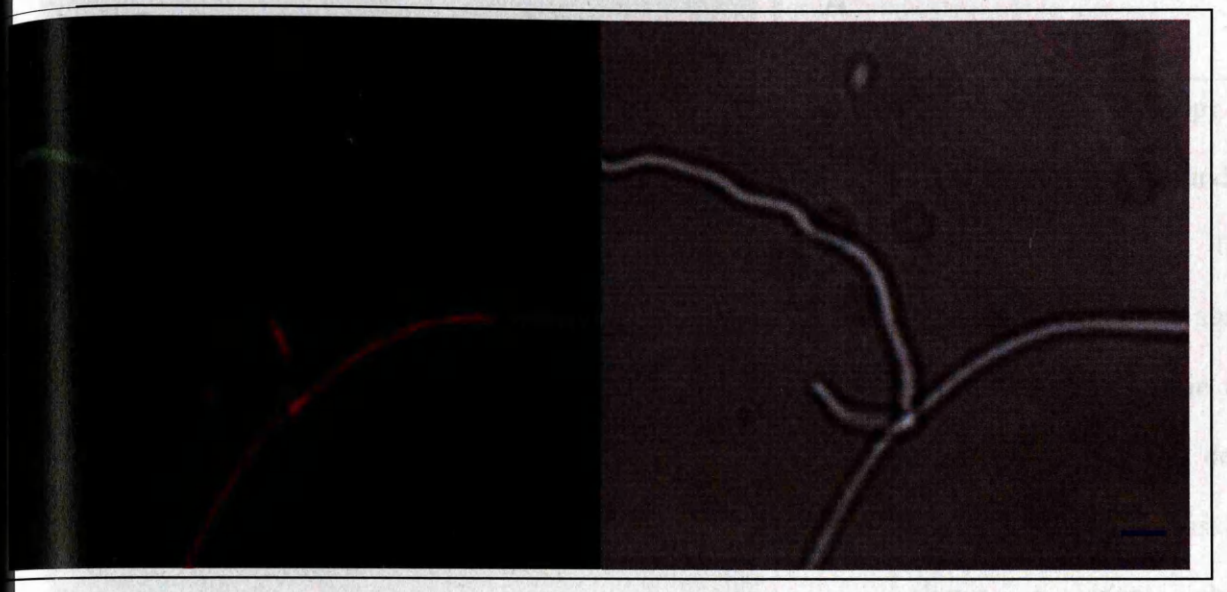
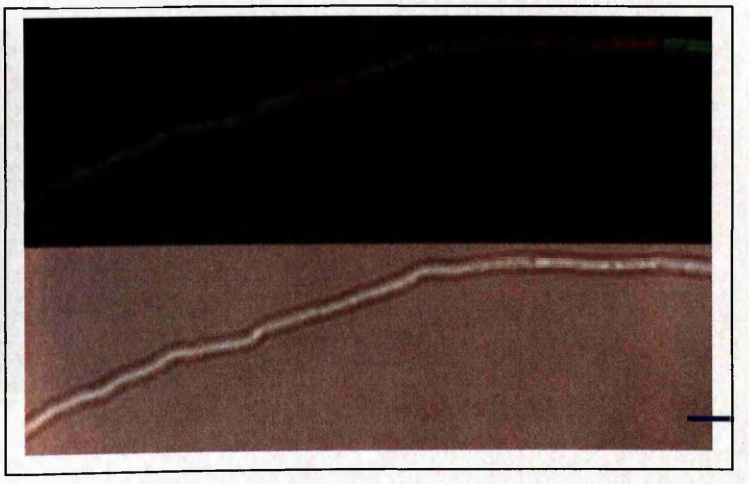


Figure 6.9: Microscopy of M145 parental strain using BacLight stain. A = substrate hyphae showing live (green) and dead (red) hyphae. B and C = spore chains demonstrating the same effect, with live / dead spores interspersed with each other. Each stained image is shown alongside an image of the same field taken using the visible spectrum. Hyphae are 72h old. Blue scale bar = 1μm.

C



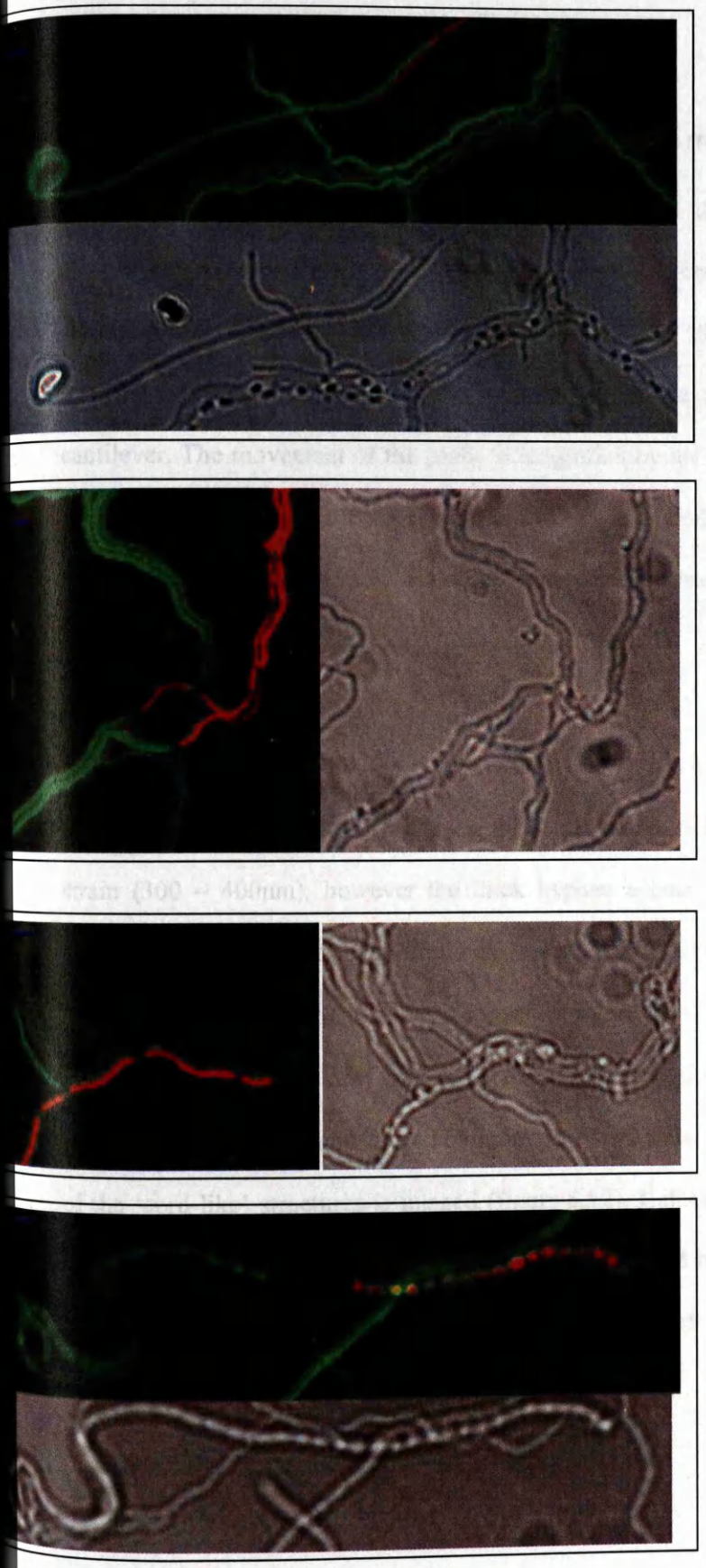


Figure 6.10: Microscopy of DSCO3843/44. A, B and C show the 'cord' like structures formed by some substrate hyphae. As in the parental M145 strain, dead and live hyphae coexist. D shows a spore chain containing both live and dead spores. As in figure 6.6, each stained image is shown with an image of the same field taken using the visible spectrum. Hyphae are 72h old. Blue scale bar = 1μm

6.3.6 Hyphal morphology under atomic force microscopy

Atomic force microscopy (AFM) allows much more detailed images of bacterial morphology to be collected. For a detailed explanation of how AFM works in a biological setting, please see the following reference (Lal and John 1994). In essence, AFM works by sensing the topography of a surface at the molecular level using a probe attached to a cantilever. The movement of the probe is magnified by the cantilever, and this is sensed by reflecting laser light off the cantilever and into a photodiode. The cantilever can also be used to 'tap' the surface, avoiding perturbation caused by dragging the tip over delicate surfaces.

AFM images of DSCO3843/44 reveal two separate types of hyphae, thin ones and thick ones (figure 6.8). The thin hyphae are the same thickness as seen in the parental M145 strain (300 – 400nm), however the thick hyphae appear to be twice as thick (600 – 800nm). There does not appear to be any correlation between the 'daughter' hyphae and their parents in terms of thickness, shown in figure 6.11. Other than the difference in thickness, there does not appear to be any other difference between individual DSCO3843/44 hyphae and those belonging to the parental M145 strain, even when one of the 'cord like' structures is imaged (figure 6.12). It does seem possible that the thick hyphae are two normal hyphae fused together, as all AFM images of them show a central spine that could be the result of improper separation of two hyphae, and this would explain the doubling in size, as well as the cords.

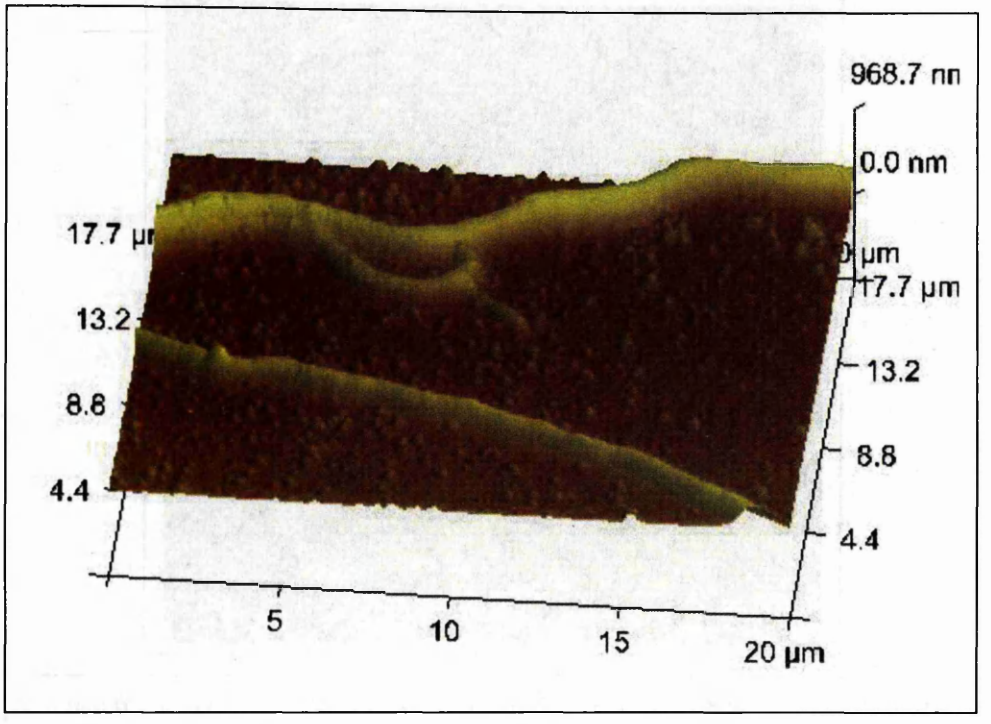


Figure 6.11: AFM image of the two kinds of hyphae found in DSCO3843/44. Uppermost hypha is an example of the thick hyphae (note the thin hypha branching from it). Lowermost is an example of the thin hyphae (identical to those found in the M145 parental strain, not shown). Scale is shown in 3 dimensions along the borders of the image. Hyphae are 72h old.

Figure 6.12: AFM image of two thick DSCO3843/44 hyphae forming part of a cord like structure. Blue scale bar represents 1µm. Hyphae are 72h old.

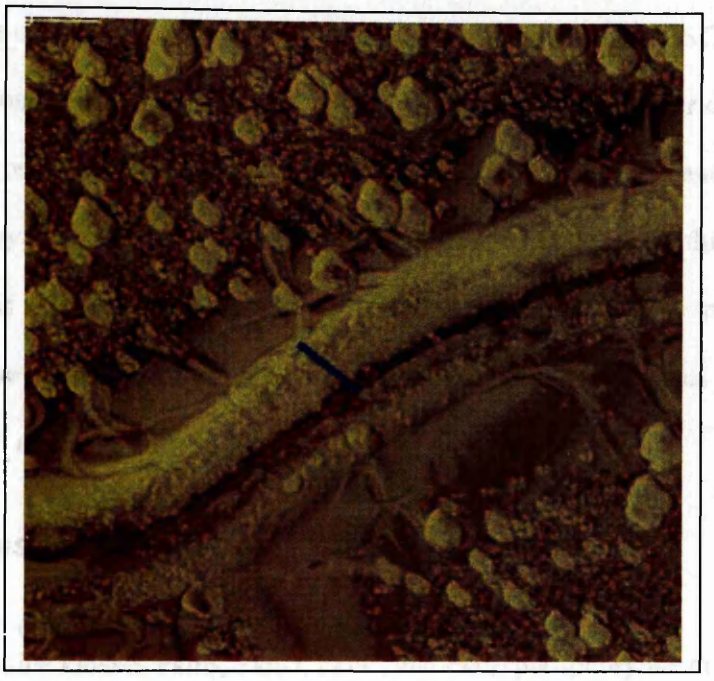


Figure 6.13: AFM image of thick DSCO3843/44 hypha Scale bar = 500nm.

6.4 Summary

SCO3843 and SCO3844 are representatives of genes coding for proteins that have been shown to be important mediators of phosphoserine / phosphothreonine signalling in both eukaryotes and prokaryotes. Thus far, bacterial FHA genes have been shown to interact with STPK molecules, however their wider function remains unknown. Construction of a double disruption, DSCO3843/44, is a step towards finding answers to this question. DSCO3843/44 has similarities and differences to DSCO3848. The similarities rest on the overproduction of antibiotic and precocious erection of aerial hyphae when plated on general rich media such as NE. The mutant also shows that combinations of phosphate and glucose, and citrate and glucose represses aerial development. This can be rescued

either by using mannitol as a carbon source, or by replacing phosphate and citrate with MES. When grown on media containing various metabolic intermediates, DSCO3843/44 reveals varied responses which depend on the intermediate supplied. Further differences are observed microscopically, as DSCO3843/44 forms 'cords' reminiscent of those formed by saprophytic fungi as well as generating a mixture of thin and thick hyphal morphologies. This reveals the possibility of SCO3843 and SCO3844 playing an important role in the regulation of the carbon flux dedicated to primary versus secondary metabolism, as well as substrate hyphal morphology.

6.5 References for Chapter 6

Grundner, C., L. M. Gay, and T. Alber. 2005. Mycobacterium tuberculosis serine/threonine kinases PknB, PknD, PknE, and PknF phosphorylate multiple FHA domains. Protein Sci **14**:1918-1921.

Kanehisa, M., S. Goto, M. Hattori, K. F. Aoki-Kinoshita, M. Itoh, S. Kawashima, T. Katayama, M. Araki, and M. Hirakawa. 2006. From genomics to chemical genomics: new developments in KEGG
10.1093/nar/gkj102. Nucl. Acids Res. **34**:D354-357.

Lal, R., and S. A. John. 1994. Biological applications of atomic force microscopy. Am J Physiol Cell Physiol **266**:C1-21.

Letunic, I., R. R. Copley, S. Schmidt, F. D. Ciccarelli, T. Doerks, J. Schultz, C. P. Ponting, and P. Bork. 2004. SMART 4.0: towards genomic data integration. Nucleic Acids Res **32 Database issue**:D142-144.

Norman E. Good, G. D. W., Wilhelmina Winter, Thomas N. Connolly, Seikichi Izawa, Raizada M. M. Singh. 1966. Hydrogen Ion Buffers for Biological Research. Biochemistry **5**:467 - 477.

7 SCO3845: An inessential phosphatase in an essential operon.

7.1 Introduction

SCO3845 codes for a ser / thr phosphatase which is conserved in all actinobacterial genomes and a large number of firmicute genomes (see chapter 3). Orthologues of this gene product have been shown to dephosphorylate the phosphorylated form of PknB orthologues (see introduction). SCO3845 lies at the start of a putative operon contained within the PknB gene cluster which includes the *ftsW* / *rodA* gene (SCO3846) and the *pbp2* or thologue (SCO3847). Disruption of SCO3846 is lethal to M145 (B. Mistry, unpublished results), thus any disruption of SCO3845 must ensure that SCO3846 is still expressed.

7.2 DSCO3845

7.2.1 Disruptant construction

To overcome the problem of keeping SCO3846 expression normal, it must be disrupted keeping the promoter activity normal. The options available at the start of this process were to either engineer a strain with the promoter located one gene downstream (i.e. at the start of SCO3846 rather than at the start of SCO3845), or to create a strain with an in – frame deletion of SCO3845 that would allow read – through to SCO3846. Analysis of the SCO3845 DNA sequence revealed 3 unique *Aat*II restriction sites within the phosphatase gene. These sites allowed a piece of the gene to be removed using restriction

enzymes, leaving the reading frame intact. This deletion was carried out in plasmid pMEF to form plasmid pMEFP- (see figure 7.1a and b). The deletion was confirmed with restriction analysis.

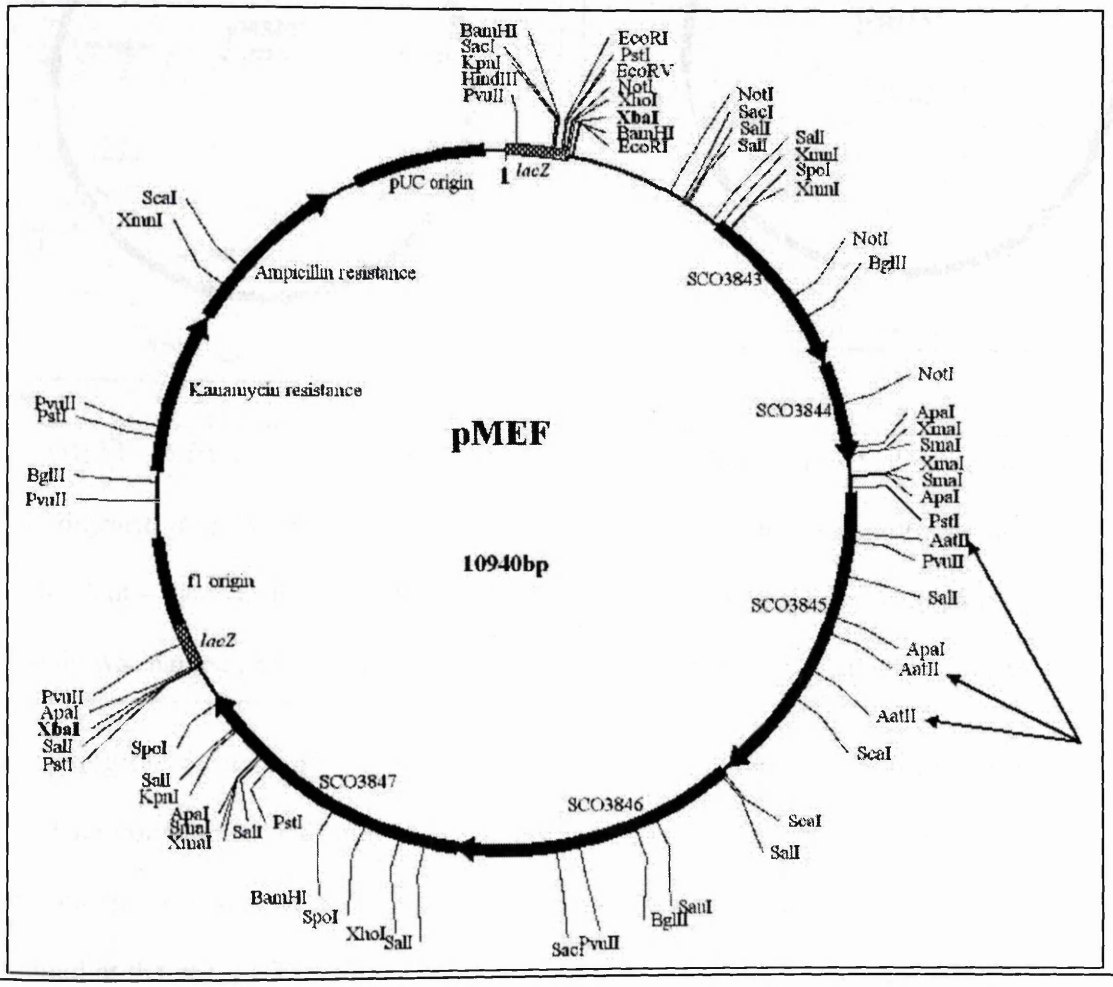
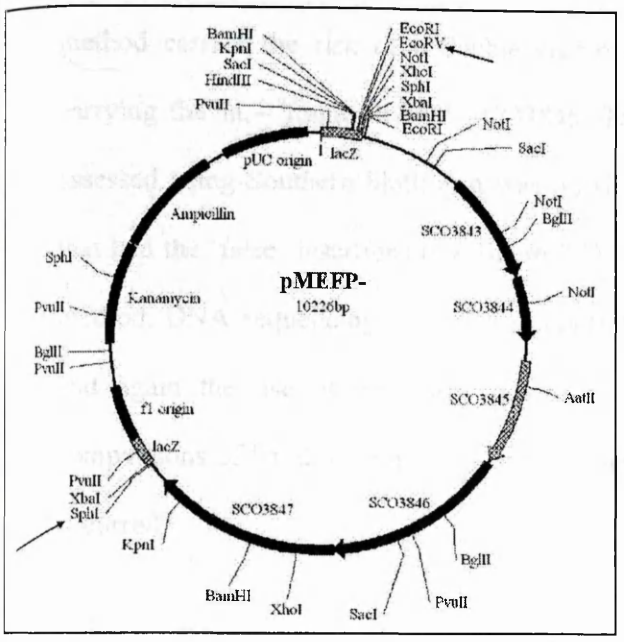
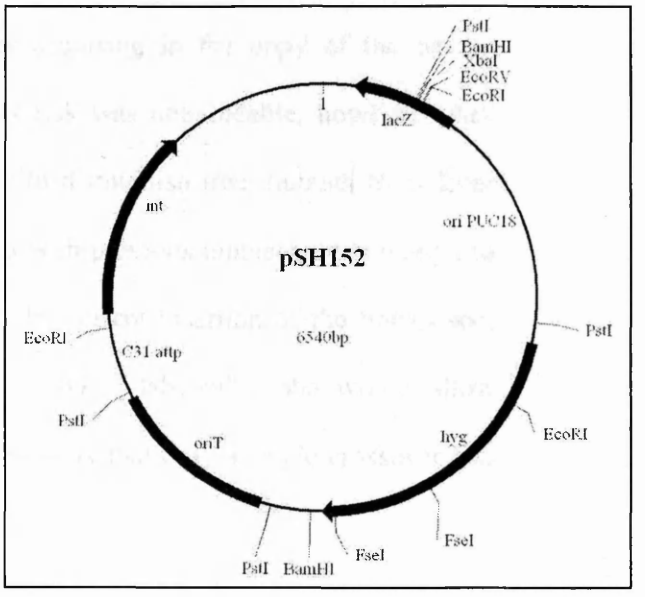


Figure 7.1a: Plasmid pMEF containing the SCO3845 operon. The unique AatII sites used to disrupt SCO3845 are indicated with arrows.

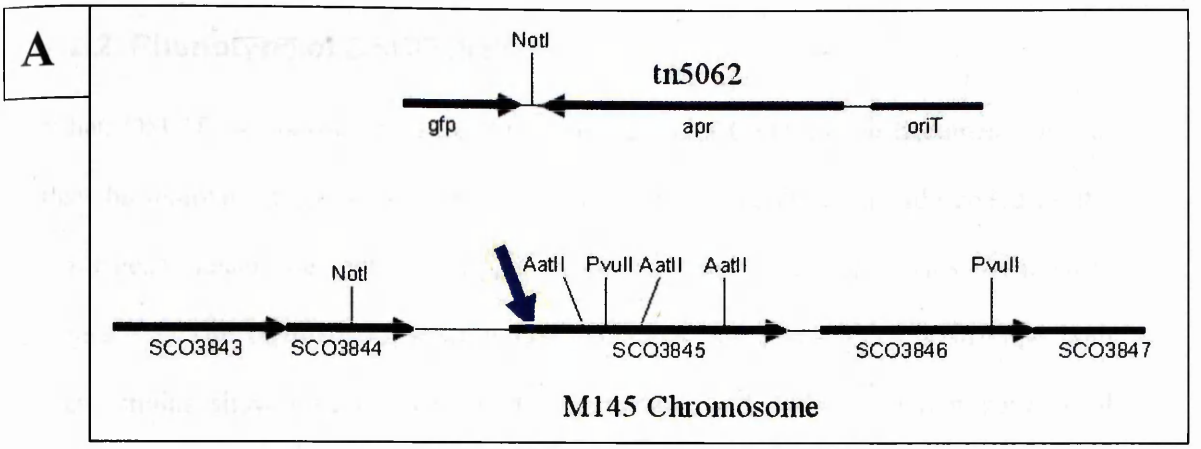


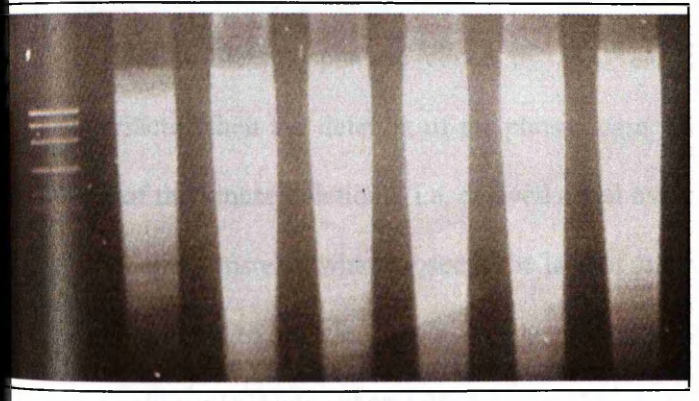


gure 7.1b: pMEFP- (left) resulting from the digestion of pMEF with *Aat*II and its subsequent re-ligation. MEFP- was digested *Eco*RV and *Sph*I, then the fragment containing the in – frame deleted copy of CO3845 was blunt – end ligated into pSH152 (which had been previously been digested *Eco*RV). The sulting plasmid was named pSHP-, and this was used in a conjugation reaction to transform M145.

Following the excision of this piece of DNA, a plasmid based on the vector pSH152 could be constructed and used in a conjugation reaction to create a strain of M145 containing two copies of the operon, one of which has a shortened copy of SCO3845 inserted at the phi – C31 attachment site in the chromosome. At the same time, another strain was constructed using plasmid pSHRC1 (kindly donated by B. Mistry) as a control. This plasmid contains an intact copy of SCO3845, and it also inserts at the phi – C31 attachment site. This strain is designated M145-RC1. The next step was to use transposon insertion to disrupt the native copy of the gene (in M145-RC1 and the potential DSCO3845) using transposon insertion SCH69.1.C06 in the standard method. This

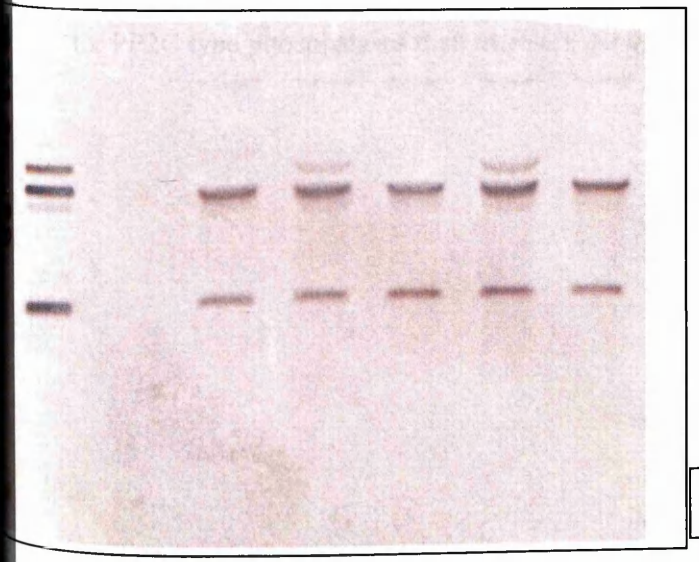
method carried the risk of a double crossover occurring in the copy of the operon carrying the in – frame deleted SCO3845. This risk was unavoidable, however when assessed using Southern blotting it was possible to distinguish true mutants from those that had the 'false' insertions (see figure 7.2). As with previous mutants made using this method, DNA sequencing was used to confirm the correct insertion of the transposon, and again the use of the entire cosmid as a DIG – labelled probe would allow comparisons of band pattern to eliminate the possibility that only a single crossover had occurred.





CH69.1.C06

M145 DSCO3845 M145 – RC1



C

Figure 7.2: Confirmation of correct insertion of Tn5062. 'A' shows the insertion site of Tn5062 and the relevant restriction sites. All strains were subsequently digested *Not*I / *Pvu*II to give the result shown on an agarose gel stained with ethidium bromide in 'B'. Subsequently, the DNA from the gel in 'B' was transferred to a membrane and Southern blotting with DIG labelled tn5062 was performed. The expected bands of 2680 and 1443 bp were detected, whilst the 4772bp fragment expected for a false insertion was not detected. The labels for each lane on blot / gel are shown between 'B' and 'C'.

7.2.2 Phenotype of DSCO3845

When DSCO3845 is compared to DSCO3848 and DSCO3843/44, it becomes apparent that the removal of the phosphatase from the putative signalling cascade coded by the *pknB* gene cluster does not have a phenotypic effect in *S. coelicolor*. This is shown in figure 7.3. The comparison between DSCO3843/44 and DSCO3848 is striking, both these strains show precocious antibiotic production, and widely different patterns of aerial hyphal development compared to the parental M145 strain. The expectation would be that if the phosphatase coded by SCO3845 and the kinase coded by SCO3848 interracted then the deletion of the phosphatase would produce a phenotype opposite to that of the kinase deletion – i.e. delayed aerial hypha development and retarded antibiotic production. Instead, what we see is the lack of any phenotype associated with its absence. This does led to questions regarding the role of SCO3845, especially as its orthologue has been shown to dephosphorylate phosphorylated PknB in *Mycobacteria* (see introduction). The answer may lie in the larger number of genes within the *S. coelicolor* genome coding for PP2C type phosphatases than there are genes coding for STPKs.

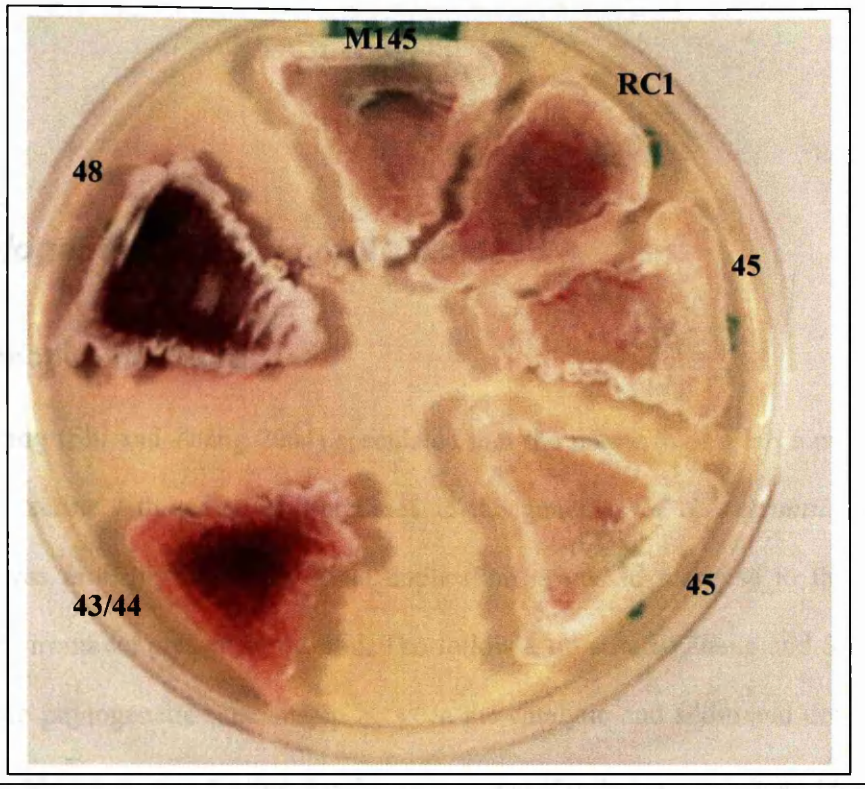
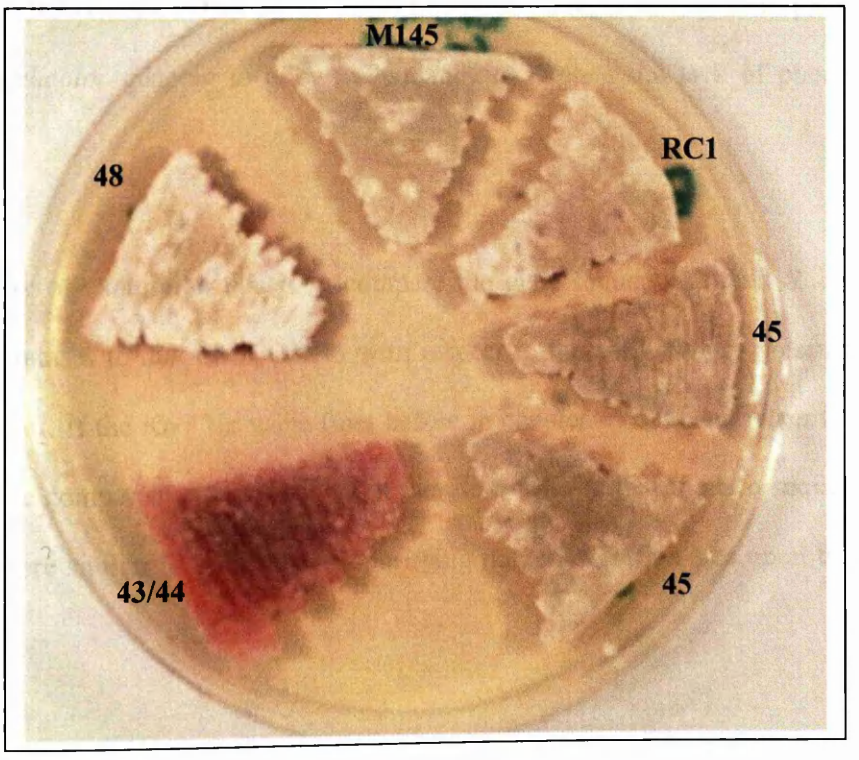


Figure 7.3: Phenotype of DSCO3845. The strains indicated are: M145 = parental M145 strain, RC1 = M145-RC1, 45 = DSCO3845, 43/44 = DSCO3843/44, 48 = DSCO3848. The image above is the result of 72h incubation at 30°C on NE buffered to pH 7.1, below is the result of 48h incubation on NMMP – G media at 30°C.



7.3 Phylogenetic analysis of SCO3845

7.3.1 Selection pressure

Shi and Zhang (Shi and Zhang 2004) speculated that the presence of a large number of 'eukaryotic' serine / threonine phosphatases in the genomes of *S. avermitilis* and *S. coelicolor* was probably due to multiple duplication events in response to the ever – changing environment present in the soil. The follow – up article (Zhang and Shi 2004) examined the phylogenetic relationship between the catalytic and additional domains. It was postulated that after horizontal transfer from a eukaryotic source, the streptomycete phosphatases diversified and recruited additional domains as signalling pathways evolved to become more specialized. The majority of *S. coelicolor* PPM phosphatases have recruited additional domains, however SCO3845 codes for a phosphatase domain only. The question thus arises: is this phosphatase sufficiently similar to any other phosphatase in the *S. coelicolor* genome to allow redundancy to explain the lack of phenotype in DSCO3845?

The best way of examining this is to compare the gene sequences of all *S. coelicolor* PPM phosphatases to that of SCO3845 with respect to the evolutionary pressure acting upon each pair. If the Ka / Ks value (met before in chapter 4) obtained when two DNA sequences are compared is less than 1.0, it therefore follows that the two sequences are becoming more similar to each other as purifying natural selection acts upon them. The

sequences compared and their subsequent Ka, Ks and Ka/Ks values are shown in table 7.1.

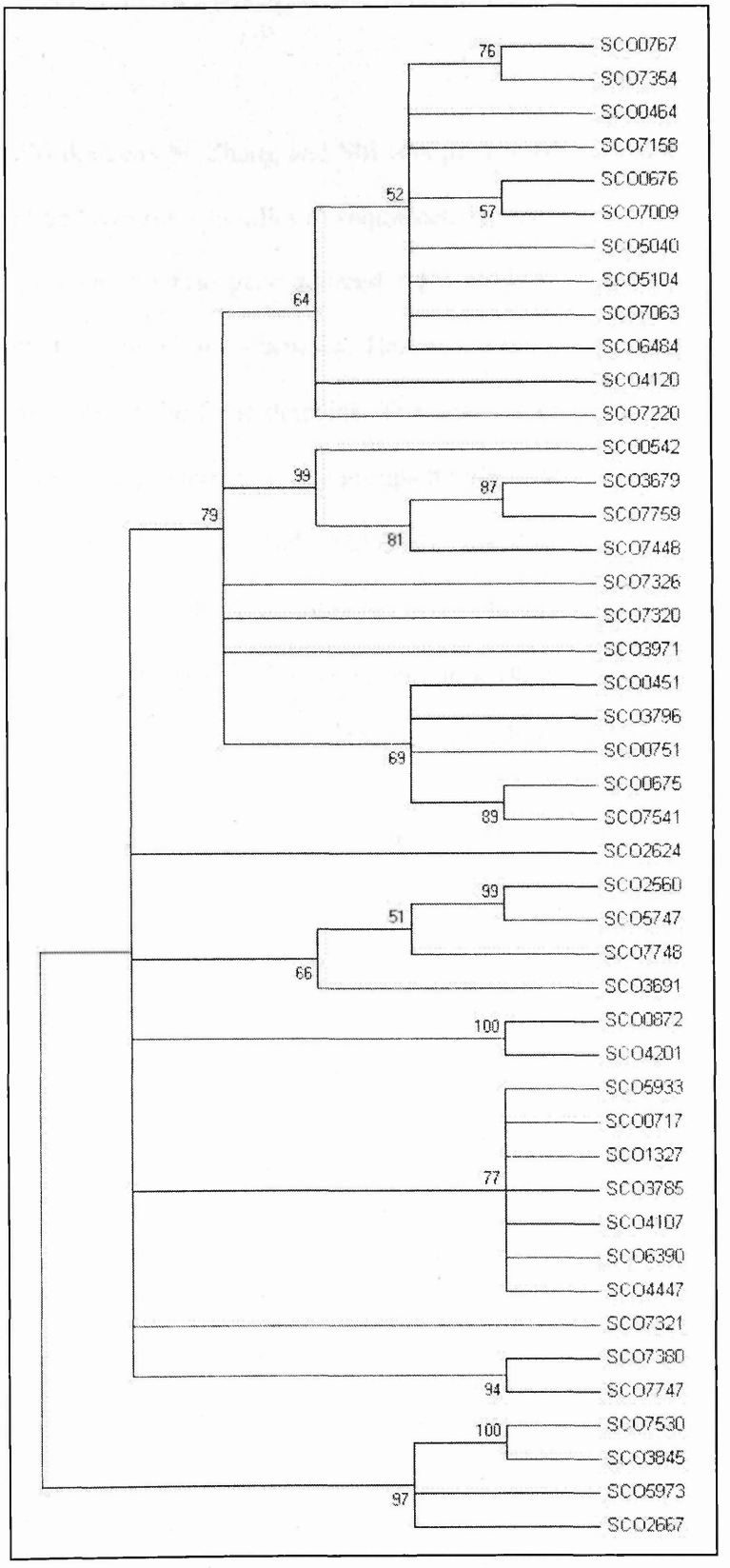
Gene	Ka	Ks	Ka/Ks
SCO7759	1.37	0.72	1.91
SCO7748	1.29	0.51	2.52
SCO7747	1.66	0.89	1.88
SCO7541	1.62	0.46	3.50
SCO7530	***0.43***	***0.69***	***0.63***
SCO7448	1.46	0.48	3.04
SCO7380	1.56	0.57	2.73
SCO7354	1.42	0.49	2.92
SCO7326	1.72	0.78	2.21
SCO7321	1.79	0.66	2.72
SCO7320	1.72	0.72	2.39
SCO7220	1.37	0.61	2.25
SCO7158	1.56	0.63	2.47
SCO7063	1.55	0.62	2.50
SCO7009	2.01	0.76	2.63
SCO6484	1.30	0.56	2.34
SCO6390	1.37	0.95	1.44
SCO5973	1.44	0.42	3.43
SCO5933	1.20	0.54	2.21
SCO5747	1.30	0.55	2.37
SCO5104	1.36	0.53	2.56
SCO5040	1.52	0.67	2.26
SCO4447	1.28	0.51	2.51
SCO4201	1.60	0.59	2.73
SCO4120	1.45	0.43	3.37
SCO4107	1.36	0.48	2.83
SCO3845	N/A	N/A	N/A
SCO3971	1.46	0.54	2.68
SCO3796	1.42	0.72	1.98
SCO3785	1.50	0.54	2.78
SCO3691	1.11	0.36	3.07
SCO3679	1.45	0.52	2.79
SCO2667	1.09	0.46	2.36
SCO2624	1.64	0.42	3.90
SCO2560	1.33	0.50	2.64
SCO1327	1.44	0.98	1.46
SCO0872	1.58	0.79	2.00
SCO0767	1.59	0.79	2.00
SCO0751	1.57	0.55	2.84
SCO0717	1.23	0.87	1.41
SCO0676	1.44	0.85	1.71
SCO0675	1.45	0.55	2.62
SCO0542	1.57	0.54	2.89
SCO0464	1.57	0.70	2.23
SCO0451	1.51	0.57	2.66

Table 7.1: Ka, Ks and Ka / Ks values for the DNA sequence coding for the SCO3845 PPM phosphatase domain when compared to similar domains from other phosphatases within the *S. coelicolor* genome. The single sequence which resulted in a Ka / Ks value less than one is indicated in bold and italicised (SCO7530).

Table 7.1 shows a quite startling degree of divergence between *S. coelicolor* PPM phosphatase domain sequences and SCO3845, with the vast majority having Ka/Ks values well above 1.0. Indeed, the job of choosing a suitably similar sequence to SCO3845 to justify the claim that it performs a similar function becomes easy, as only a single sequence falls into the category of being under purifying selection, SCO7530.

7.3.2 Phylogeny

If two sequences are evolving according to purifying selection, the reasoning follows that they are likely to be close to each other on a phylogenetic tree. This turns out to be the case for SCO3845 and SCO7530 (fig 7.3), however this relationship only appears to be true for genes diverging from the same node – the next closest sequences in terms of phylogeny have Ka / Ks values greater than one (SCO2667 and SCO5973).



ure 7.4: Phylogenetic tree showing tionships between *S. coelicolor* PPM sphatase domains. The tree was structed using PhyML and the numbers at es are percentage values from 500 strap replicates. Only nodes supported 0% or greater bootstraps were retained. tree was drawn using MEGA3.1. All ysis was carried out on DNA sequences.

The previous analysis of *Streptomyces* PPM domains by Zhang and Shi was performed using amino acid sequences, and they identified two sub - families of sequences that they surmised had resulted from duplication of a phosphatase gene sourced from an early eukaryote. SCO3845 is contained in the smaller of these sub – families. The second sub – family is much larger, and contains the majority of the PPM domains. The analysis of DNA sequences here allows this sub – family to be grouped into five groups by virtue of bootstrap values greater than 50%, with SCO2624 and SCO7321 forming phylogenetically distinct branches on their own. Similarly, the smaller sub – family containing SCO3845 can be split, with SCO3845 and SCO7530 residing within group 6 (see figure 7.5). This analysis would allow a more detailed examination of the role of each of the *S. coelicolor* PPM phosphatases, however it is outside the subject of this thesis to do so.

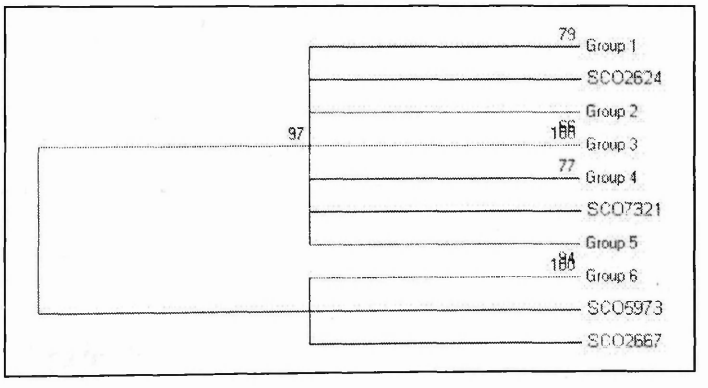


Figure 7.5: Condensed version of figure 7.4, showing the major groups of PPM domains discovered through phylogenetic analysis. The sub – families mentioned in the text branch from the first node (97% bootstrap support).

7.4 Conclusion

SCO3845 may well be a redundant gene that has been preserved in the *S. coelicolor* genome due to its position upstream of the essential SCO3846 gene in the same operon. Although its genetic position did not allow a mutant to be made in as straightforward a manner as usual, it was possible to create one by taking advantage of restriction sites that allowed an in – frame deletion to be constructed. The resulting disruption of SCO3845 did not appear to affect the phenotype in a significant manner compared to the parental M145 strain, although it may be that different growth conditions are required for a DSCO3845 phenotype to arise. This seemed to confirm the theory that SCO3845 is indeed a redundant gene, however it would be advantageous to identify the gene(s) that were able to supplant its function. Analysis of selection pressure acting on all *S. coelicolor* PPM catalytic domains revealed that SCO7530 is the most likely candidate, as it is the only gene found to be evolving in the same way as SCO3845. This is supported by phylogenetic analysis, which itself supported previous work by Zhang and Shi, and further identified increased complexity in relationships between catalytic domains than said previous analysis had revealed.

7.5 Chapter 5 references

Shi, L., and W. Zhang. 2004. Comparative analysis of eukaryotic-type protein phosphatases in two streptomycete genomes 10.1099/mic.0.27057-0. Microbiology **150**:2247-2256.

Zhang, W., and L. Shi. 2004. Evolution of the PPM-family protein phosphatases in Streptomyces: duplication of catalytic domain and lateral recruitment of additional sensory domains 10.1099/mic.0.27480-0. Microbiology **150**:4189-4197.

8 Properties and Targets

8.1 *Introduction*

Previous work on PknB like proteins has been heavily biased toward analysis of expressed protein from *E. coli*. The analysis has revealed that PknB – like proteins autophosphorylate, and that they can subsequently be dephosphorylated by the cognate phosphatase present in the *pknB* gene cluster (see Introduction). It has also been shown that PknB like proteins form a dimer (Wehenkel et al.; Ortiz-Lombardia et al. 2003; Young et al. 2003) when complexed with ATP. The substrates of PknB like proteins have been detailed elsewhere in this work (see Introduction), and their structure is shown in Figure 4.2. The purpose of studying SCO3848 from a proteomic perspective is to confirm that it shares the same properties as those orthologues that have been studied elsewhere, and thus to justify extrapolations from data obtained from proteins in different organisms. It is also the intention to identify potential targets through the analysis of the proteome of disruptant strains.

8.2 *Cloning and overexpression of SCO3848 in E. coli*

8.2.1 Site directe d mutagenesis and plasmid construction

Previous work on SCO3848 had led to the creation of plasmid pLK1 (a derivative of pIJ2925), containing SCO3848, SCO3849 and a part of SCO3850. This plasmid was used in a site directed mutagenesis reaction (see Materials and Methods) to introduce a *Nde*I restriction site at the start of SCO3848, retaining the 'ATG' start codon (named plasmid pHit-K). It was not deemed necessary to remove SCO3849 and the fragment of SCO3850 as they are transcribed from the opposite DNA strand to SCO3848, and there is a very strong terminator in the form of a stem – loop structure

between SCO3848 and SCO3849. Overexpression in *E. coli* requires an inducible promoter, and such a promoter (P*lac*Z) is present in the 'pET' system from Promega. The expression fragment was cloned from pHit-K and into a pET16b derivative (kanamycin resistance in addition to ampicillin resistance) using restriction enzymes *Nde*I and *Bam*HI (see Figure 8.1). The resulting plasmid (pETK1) was used to transform JM109 *E. coli* cells, and a stock of the plasmid was made with which to transform other strains which would be used to express SCO3848.

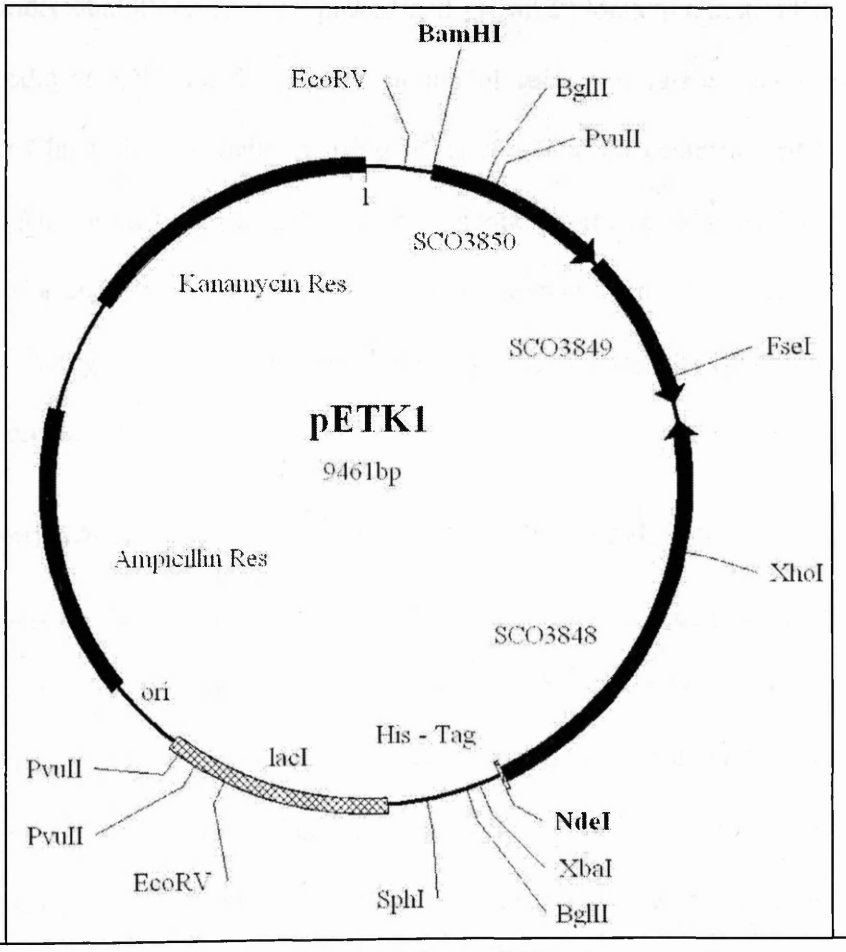


Figure 8.1: pETK1, a pET16b derivative containing a DNA fragment for expression of full length *S. coelicolor* PknB coded for by SCO3848. The enzyme sites used to insert the fragment into the vector, *Nde*I and *Bam*HI, are highlighted in bold.

8.2.2 Ex pression in *E. coli*

There are several *E. coli* strains available which have different properties suitable for different expression systems. The strains chosen were BL21 (*lon* and *ompT* protease deficient) and 'Novablue' (Tet^r, *recA, endA*). pETK1 was introduced into each cell line via electroporation. After plate selection on LB media supplemented with kanamycin, individual colonies were picked and grown in kanamycin supplemented liquid LB media at 37°C for 6 hours. Induction of cells was carried out in liquid culture as per Materials and Methods using IPTG at a final concentration of 0.1mM for one hour. After induction, the cells were resuspended in the appropriate buffer and sonicated until a clear lysate was obtained. This lysate was then centrifuged using a bench top centrifuge to remove the insoluble fraction, leaving the membrane and soluble fractions in solution.

8.2.3 In vestigation into the localisation of SCO3848 in E. coli

Web – based structure and localisation prediction programmes such as HMMTOP (Tusnady and Simon 2001) predict that SCO3848 has a single transmembrane domain predominantly consisting of hydrophobic amino acids, and thus it is likely to localise to membranes. If this is the case, then SCO3848 will be predominantly found in the membrane fraction of *E. coli* cell homogenates. The membranes can be extracted from 'crude soluble' fractions (supernatant from bench – top centrifugation of whole cell homogenate) using ultracentrifugation at 150000 rpm using an SW60Ti rotor for 1hr. Subsequent western blotting of the membrane fraction using anti penta – his antibodies revealed that overexpressed SCO3848 was predominantly found in the membrane fraction from *E. coli* (Figure 8.2).

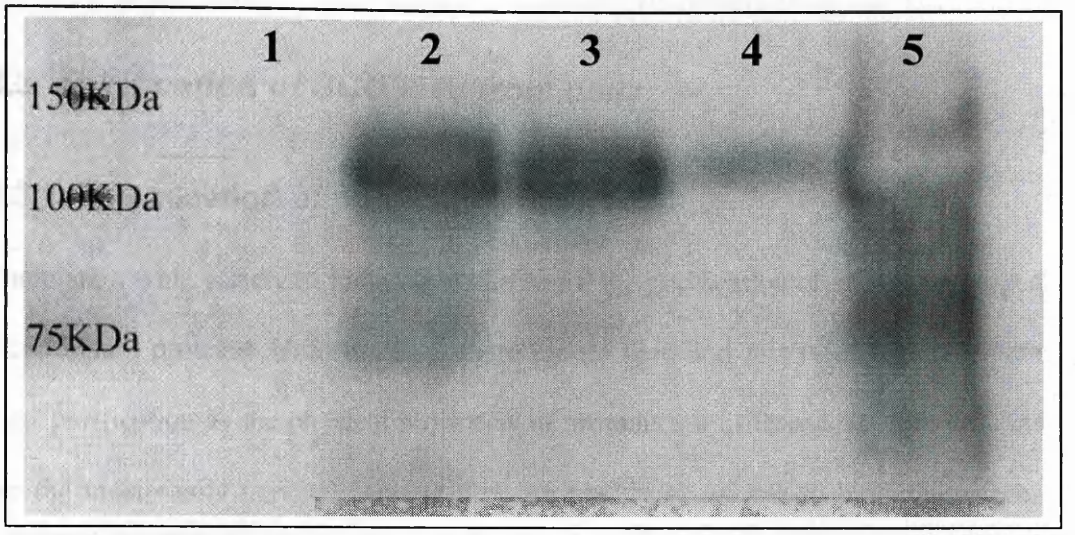


Figure 8.2: Western blot of *E. coli* cellular fractions after overexpression of SCO3848. Antibody used is HRP conjugated anti penta – His from Invitrogen. Lane 1 = whole cell extract (WCE) from uninduced *E. coli* transformed with pLK1. Lane 2 = WCE from induced cells containing pLK1. Lane 3 = insoluble fraction collected after centrifugation of WCE using bench top centrifuge. 4 = soluble fraction. 5 = membrane fraction. All samples were suspended in sonication buffer (see Materials and Methods). Protein concentration was determined using the 'Bradford' method, and equal amounts of total protein were added to each lane.

8.3 *Purification of SCO3848 from cells*

8.3.1 Opti misation of purification methods

There are a wide variety of methods available for the purification of poly – his tagged recombinant proteins. Unfortunately, there are no hard and fast rules which govern their purification as the physical properties of proteins are different. Soluble proteins are the most easily purified because they are readily accessible to binding surfaces due to their already being in solution. Some proteins form 'inclusion bodies', insoluble aggregates of misfolded proteins that can be difficult to purify without solubilisation (Carrio and Villaverde 2002; Villaverde and Carrio 2003; Villaverde and Mar CarriÃ³ 2003). In addition, expression of membrane – bound proteins presents another challenge as they are much more likely to form inclusion bodies. Solubilisation of inclusion bodies can denature proteins, and thus studying their function requires refolding. One method for avoiding the need for denaturation of proteins is to express only the domains of interest as soluble proteins, and this method has been used extensively in the study of transmembrane STPKs (Av-Gay, Jamil, and Drews 1999; Boitel et al. 2003; Cowley et al. 2004; Jin and Pancholi 2006). Another method is to attempt to remove the protein from the membrane under denaturing conditions. A detergent is needed for the removal of the membrane, however detergent choice must take into account the methods available for processing downstream, as some purification methods are incompatible with certain detergents. Three detergents (CHAPS, SDS and TritonX-100) were tested for their ability to provide a good yield of protein from *E. coli* membranes after overexpression of SCO3848 (Figure 8.3).

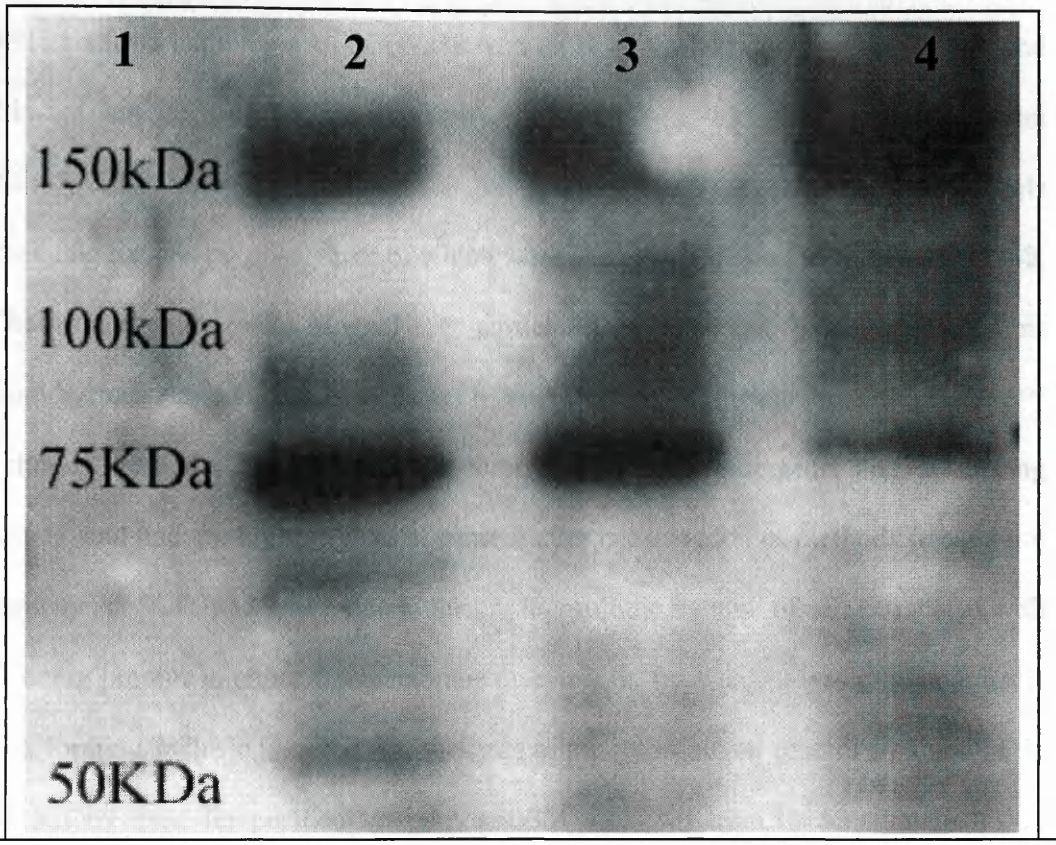


Figure 8.3: Western blot of *E. coli* BL21/pETK1 whole cell extract (WCE) prepared after induction as described in the text. The primary antibody used in this blot is described in section 8.4.1. Lane 1 = Molecular weight marker, Lane 2 = WCE in sonication buffer containing 1% Triton X-100, Lane 3 = WCE in sonication buffer containing 1% SDS, Lane 4 = WCE in sonication buffer containing 50mM CHAPS.

pETk1 allows expression of SCO3848 with an N – terminal deca – histidine tag, and this tag can be used for ion – affinity purification (as histidine residues have an affinity for divalent cations such as Ni^{2+}). There are many commercial products available for this purpose, three of which were used in an attempt to purify SCO3848. The products tried were MagneHis™ particles from Promega, Vivaspin™ columns from Sartorius and HiTrap™ columns from Amersham. Of these three, the HiTrap™ columns were the most versatile in terms of use with detergents and denaturing agents, and had the highest yield of protein after optimisation of methods (results not shown). As SCO3848 was found in the crude insoluble fraction of cell extracts as well as being present in either the membrane or cytosolic fractions, it was assumed that it was forming inclusion bodies. To purify as much recombinant protein as possible, in all buffers used for purification purposes 8M urea was used to denature inclusion bodies and 2% w/v Triton-X 100 detergent was included to disrupt membranes. The HiTrap™ system requires all buffers to contain a quantity of imidazole to decrease the amount of non-specific binding of *E. coli* proteins to the matrix. There is an optimum imidazole concentration for both wash and elution buffers that allows purification of the recombinant protein with a minimum amount of contamination from host proteins. In the case of SCO3848, a wash buffer concentration of 10mM and an elution buffer concentration of 300mM imidazole was found to be optimum (Figure 8.4). Only the 130KDa band was evident in the purified samples, perhaps reflecting a difference in conformation that exposed the histidine tag.

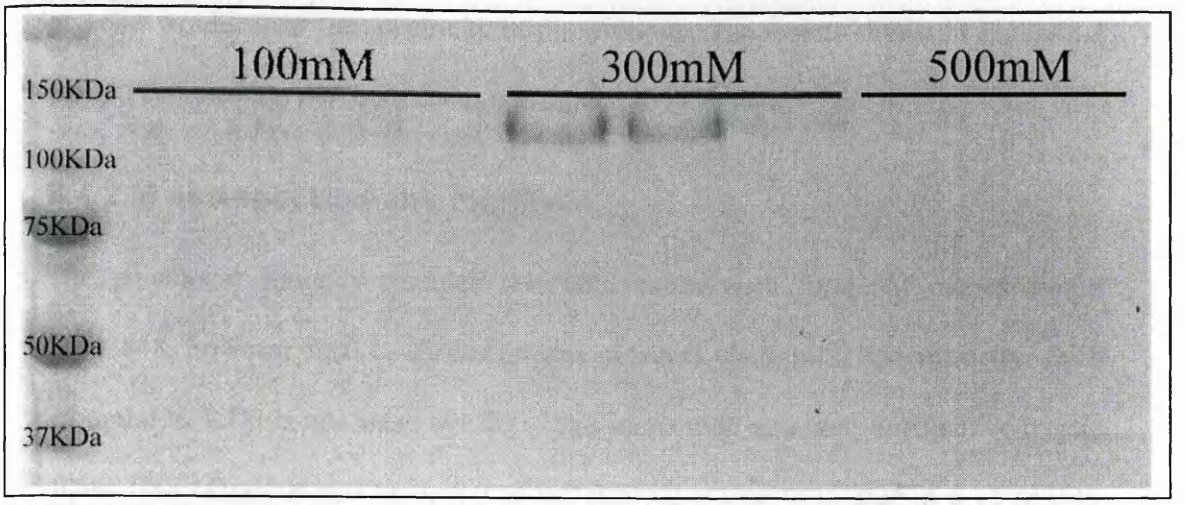


Figure 8.4: Optimisation of purification conditions. The gel shown has been stained with Coomassie blue. The molecular weight markers are indicated on the left. Imidazole concentration is indicated above groups of 3 lanes, each group of three represents sequential fractions viewed from left to right. The gel is an 8% polyacrylamide gel, samples were boiled prior to loading using a heat block.

8.4 *Confirmation of band identity*

8.4.1 Pol yclonal antibody method

Previous results have shown that when run on a gel, full length SCO3848 migrates as two bands, one at approximately 75kDa (the predicted size of SCO3848 is 72kDa), the other at approximately 130kDa. Both of these bands has been seen on western blots using anti penta – his antibodies conjugated to horseradish peroxidase, which suggests that both bands correspond to a his – tagged protein. However, there was a possibility that one or both of these bands represented a false – positive result. To check this scenario, a polyclonal antibody (raised in mice against a custom oligopeptide) specific for a portion of SCO3848 was obtained from Antibody Core of Southwestern Medical Centre, Dallas, Texas. This antibody was used in a western blot of total cell extract, as the degree of non – specific binding to native *E. coli*

proteins would show the specificity of the antibody. The results shown in Figure 8.3 help to demonstrate that both bands correspond to SCO3848.

8.4.2 M ass spectrometry method

The polyclonal antibody revealed that both bands were forms of overexpressed SCO3848, however final confirmation was obtained using mass spectrometry. Both 75 and 130 KDa bands were cut from polyacrylamide gels and purified. A tryptic digest (see Materials and Methods) of the purified protein was used in a LCMSMS mass spectrometer, and the resulting fragments were compared to the *S. coelicolor* and *E. coli* proteome sequences for fragments of the same size. All fragments from the digest matched fragments from SCO3848, with similar peptides detected for each band (data provided in appendix to this chapter).

8.5 *Investigation of SCO3848 phosphorylation*

Studies of PknB orthologues have consistently revealed the ability of these proteins to autophosphorylate when overexpressed in *E. coli* (Boitel et al. 2003; Chopra et al. 2003; Dasgupta et al. 2006). It is reasonable to assume that if PknB from *S. coelicolor* has similar properties in terms of phosphorylation, then it is valid to speculate that is shares further similarities with its orthologues. To check this, the purified recombinant protein was stained using the Pro-Q® Diamond stain from Molecular Probes. This stain exclusively binds to proteins phosphorylated on serine, tyrosine or threonine residues. This offers a simple method of checking whether *S. coelicolor* PknB expressed in *E. coli* autophosphorylates in the same way as the mycobacterial orthologues already described (see Introduction). In Figure 8.5, the purified *S. coelicolor* PknB is shown fluorescing whilst the control (unphosphorylated bovine serum albumin) remains dark, although the band is very obviously present when

viewed on the Coomassie stain. This confirms that *S. coelicolor* PknB autophosphorylates in *E. coli* on tyrosine, threonine or serine residues. Again, only the band that migrates at the 130KDa region has stained, which may provide a clue to the functional state of this protein.

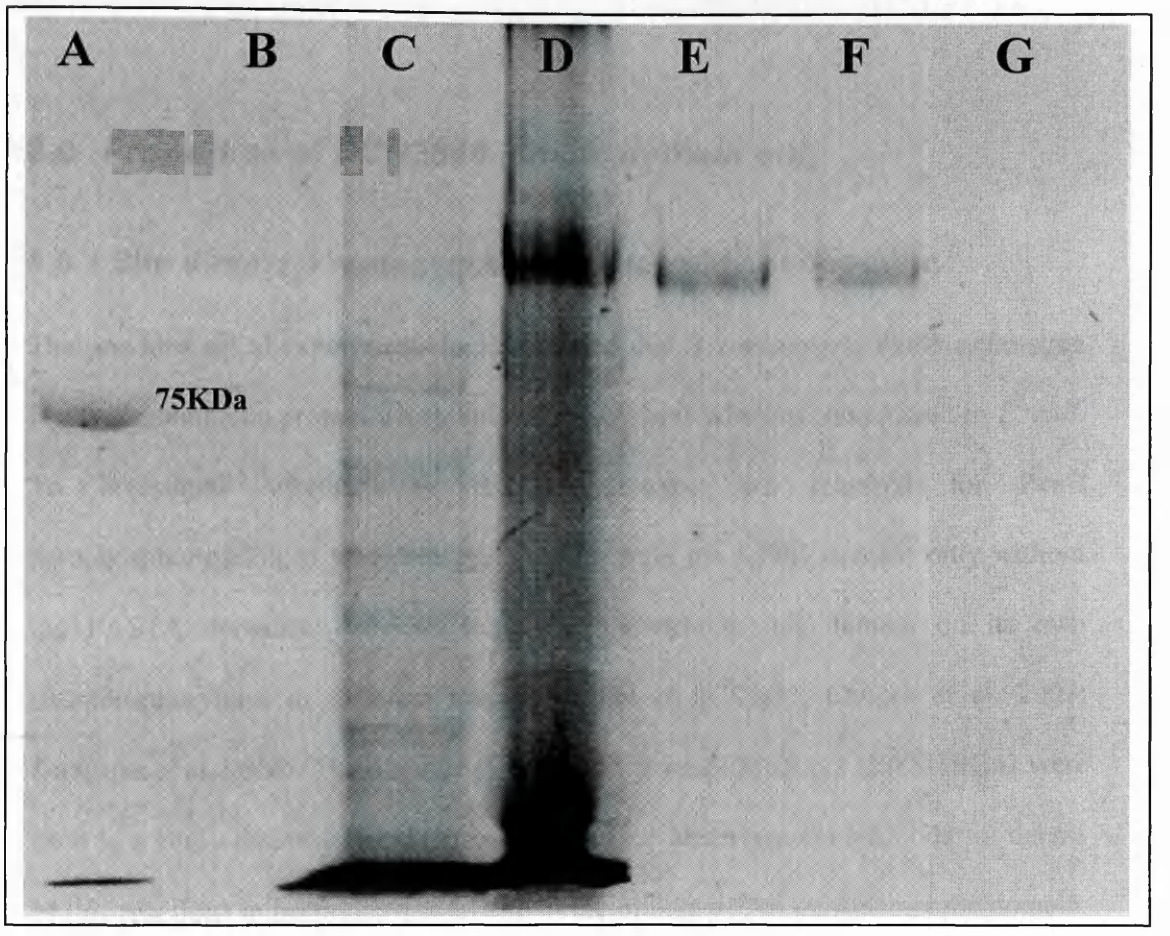


Figure 8.5a: Purified *S. coelicolor* PknB stained with the Pro-Q® Diamond stain. Lanes are as follows: A = Molecular weight marker (75KDa band has stained); B = BSA as negative control; C = WCE from *E. coli* cells before induction with IPTG; D = wash fraction from purification steps; E to G = purified *S. coelicolor* PknB.

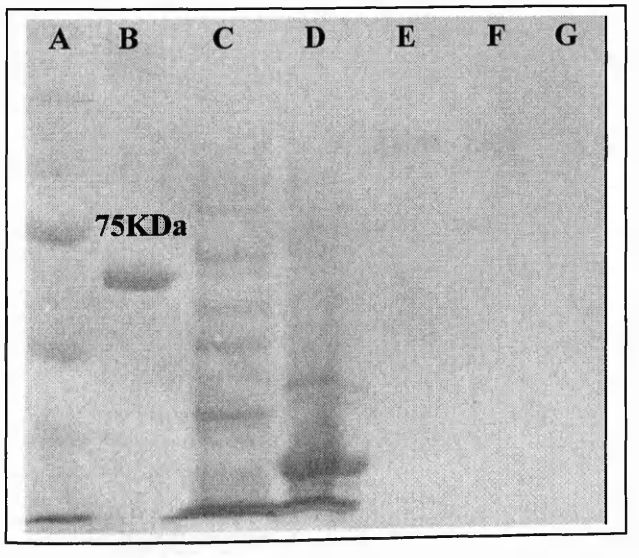


Figure 8.5b: Coomassie stain of gel shown in 8.5a. Note the presence of BSA in lane B.

8.6 *Properties of SCO3848 kinase domain only*

8.6.1 Site directe d mutagenesis and plasmid construction

The previous set of experiments had confirmed that *S. coelicolor*'s PknB orthologue is a transmembrane protein which autophosphorylates when overexpressed in *E. coli*. To investigate whether the PASTA domains are required for PknB autophosphorylation, it was decided to overexpress the STPK domain only without the PASTA domains. Previous work has shown that this domain on its own autophosphorylates in different species (Boitel et al. 2003; Chopra et al. 2003; Dasgupta et al. 2006). The plasmid pHit – K and primers MidKin1 and MidKin2 were used in a site – directed mutagenesis reaction (see Materials and Methods for details of this reaction) to introduce a *Nde*I site in the middle of the transmembrane domain. Following digestion of the resulting plasmid (pHit – K2) using *Nde*I and purification of the excised ~1kb band from an agarose gel, the STPK domain was ligated into pET16b which had been digested *Nde*I to form plasmid pETK2 (Figure 8.6). Orientation of the kinase domain was confirmed using sequencing (results not shown).

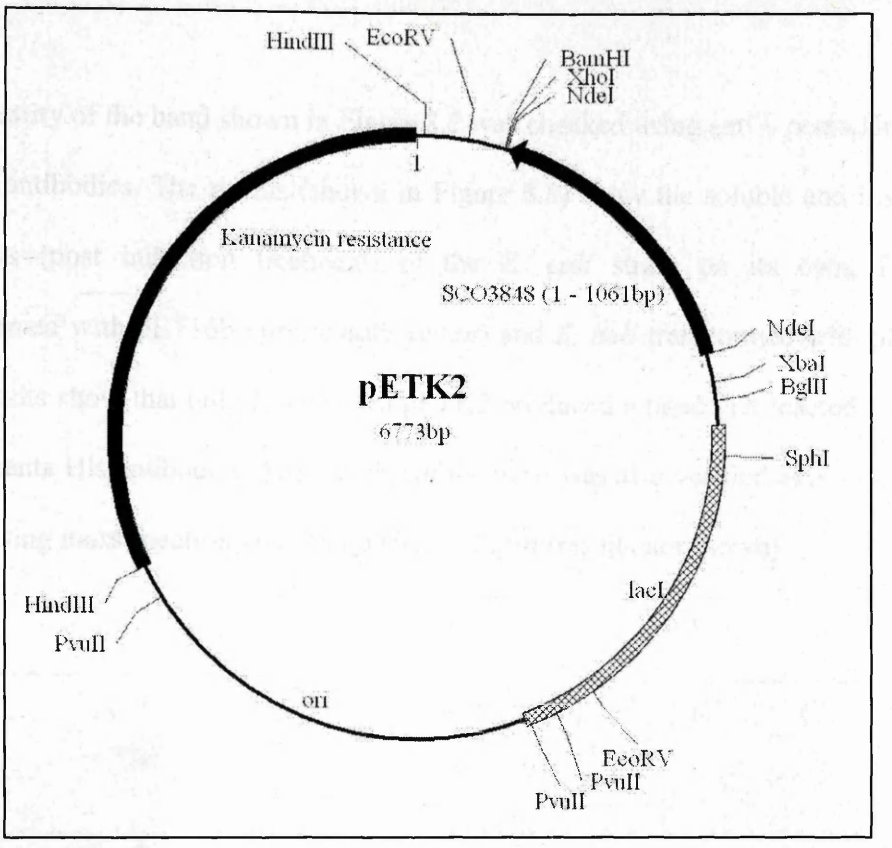


Figure 8.6: pETK2, used for overexpression of *S. coelicolor* PknB.

8.6.2 O verexpression in *E. coli* and purification

Identical conditions to those used for overexpression of full – length *S. coelicolor* PknB (see Materials and Methods and section 8.2.2) were used to overexpress the STPK domain only after pETK2 had been transformed into Novablue® *E. coli*. As before, the HiTrap™ system was used for purification of the expressed protein (see Figure 8.7).

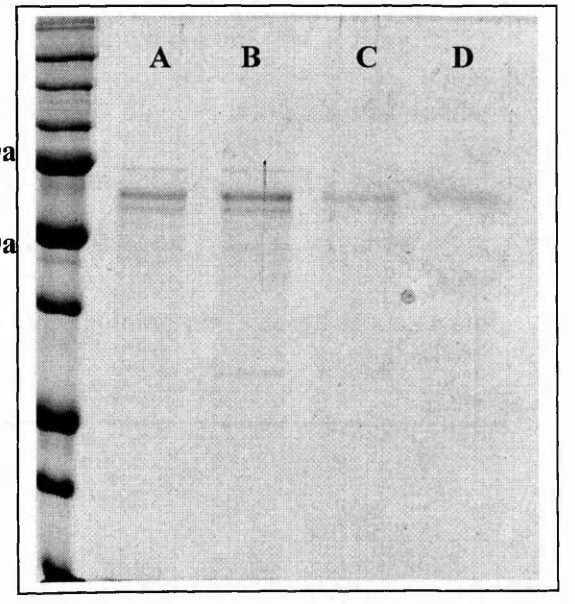


Figure 8.7: Purification of the STPK domain of *S. coelicolor* PknB. Lanes A – C = Purified domain sequentially eluted in buffer containing 300mM imidazole. Lane D = Purified protein sequentially eluted in buffer containing 500mM imidazole. Gel is silver stained.

The identity of the band shown in Figure 8.7 was checked using anti – penta His HRP tagged antibodies. The results (shown in Figure 8.8) show the soluble and insoluble fractions (post induction treatment) of the *E. coli* strain on its own, *E. coli* transformed with pET16b only (empty vector) and *E. coli* transformed with pETK2. The results show that only *E. coli* with pETK2 produced a band that reacted with the anti – penta His antibodies. The identity of the band was also verified as *S. coelicolor* PknB using mass spectrometry after a tryptic digest (results not shown).

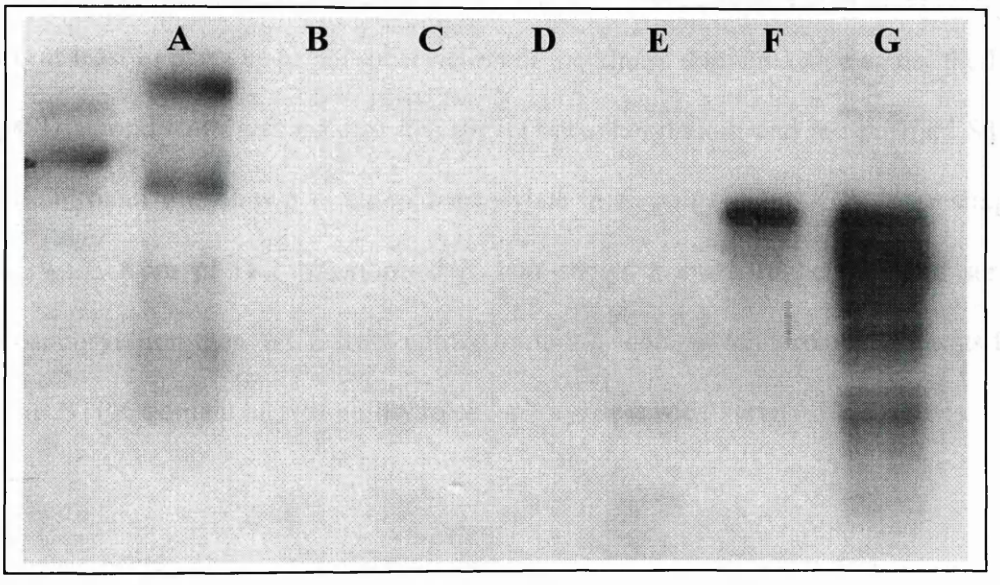


Figure 8.8: Western blot of STPK domain using penta – His antibodies tagged with HRP. Unlabelled lane = molecular size marker, A = positive control (WCE from BL21 expressing whole *S. coelicolor* PknB from pETK1), B & C = insoluble and solube fractions of Novablue® *E. coli* with no additional plasmid. Lanes D & E = soluble and insoluble fractions of Novablue® cells transformed with pET16b. Lanes F & G = soluble and insoluble fractions of Novablue® cells containing pETK2. All strains were induced with IPTG.

The size of the STPK domain bands detected in figure 8.8 are again larger than expected, being roughly double the expected size of 38KDa. Again this discrepancy may contain a clue to the function of the protein.

8.6.3 C onfirmation of phosphorylation

Previous work has shown that mycobacterial orthologues of PknB autophosphorylates when overexpressed in *E. coli*. The same has been shown for *S. coelicolor* PknB, (section 8.5). It was prudent to check the same for the kinase domain purified in the previous section to check whether the phosphorylation demonstrated in section 8.5 was at least in part due to phosphorylation of the kinase domain. Once again, the Pro-Q® Diamond stain was used to investigate phosphorylation, and the purified STPK domain band was shown to autophosphorylate in *E. coli* (Figure 8.9). Interestingly, the WCE from pETK2 transformed *E. coli* shows a much higher level of ser/thr phosphorylation than WCE from untransformed *E. coli*, which strongly suggests that PknB STPK domain has the ability to phosphorylate a wide variety of substrates.

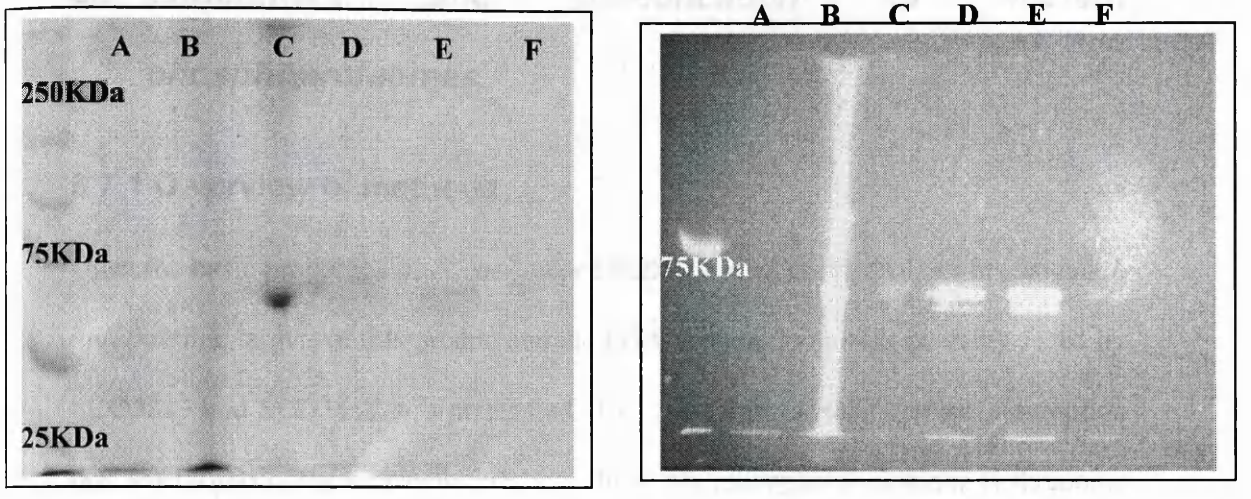


Figure 8.9: Confirmation of autophosphorylation of kinase domain. The image to the left is coomassie blue stained, the image to the right is an image of the same gel stained with the Diamond Q phosphoprotein stain. Lane key: A = WCE from untransformed induced *E. coli*, B = WCE from *E. coli* transformed with pETK2 induced with IPTG, C = BSA negative control, D – E = purified kinase domain.

8.7 *Enrichment and identification of mutant phosphoproteomes.*

8.7.1 O verview of methods

Once the basic properties of *S. coelicolor* PknB had been confirmed, an investigation into possible targets of this protein and the FHA domain containing proteins coded by SCO3843 and SCO3844 was performed. This investigation relies on the assumption that any targets of these proteins are likely to be phosphorylated on serine or threonine residues. Assuming this is true, a commercial column is available that can select only proteins phosphorylated on these residues. Once these proteins have been isolated from the WCE, mass spectrometry can be used to identify the proteins isolated. Differences in protein expression between DSCO3848, DSCO3843/44 and the parental M145 strain were identified.

8.7.2 Enrichment of phosphoproteomes

Proteomics has allowed entire proteomes to be separated out on gels and examined by comparison of different patterns of protein expression. The limitations of this system include that proteins which have low levels of expression are missed by current staining methods, and that membrane proteins are not well resolved (Garry L. Corthals 2000; Gygi et al. 2000). This has led to the development of gel – free methods that rely on mass spectrometry to identify proteins (Macek et al. 2007). This is only possible due to the availability of genomic sequence data, as peptide fragments identified are compared to possible fragments from the sequence data. Again, it is possible that proteins expressed at low levels will be lost as their signals are masked

by more abundant species. In the case of DSCO3848 and DSCO3843/44, it is very likely that the targets of the STPK and FHA domains will be differently phosphorylated in the mutants than in the M145 parental strain. As the sub-set of proteins in the *S. coelicolor* proteome which are phosphorylated are the only proteins of interest to us, it was decided to pre-fractionate whole cell extract to isolate phosphorylated proteins only. The method used was the method described in section 2.8.4 of Materials and Methods using the PhosphoProtein purification kit from Qiagen, followed by concentration of the isolated phosphoproteins using Nanosep columns.

8.7.3 Gel fractionation and LC/MS/MS comparison of mutant and parental phosphoproteomes

Following the isolation and concentration of the phosphoproteome, further pre-fractionation was performed to aid the detection of low abundance proteins. The use of 1-D gel electrophoresis allowed a degree of analysis of the protein expression profile, as well as affording a simple method of fractionating the proteins by their size. This method allowed the entire phosphoproteome to be analysed by digestion of an entire lane of a gel using trypsin, and it also allowed bands of interest to be isolated. Figure 8.10 shows the 1-D gel used for pre-fractionation.

The experiment outlined in this section (8.7) is of a preliminary nature, and as such the results need to be confirmed through methods other than those described here. It is the author's opinion that it is worthwhile to present these results, as clues to function of the deleted genes will be contained within them. Figure 8.10 shows the band pattern from the second of two replicates of this experiment, the first having presented a similar band pattern with similar peptide fragments identified by LC-MS/MS. It

should also be noted that several of the potentially phosphorylated proteins identified within the following tables are predicted to be extracellular or secreted, and as such are not likely to be phosphorylated. This is likely to be due to non – specific binding in the Quiagen column. Phosphorylation of these proteins could be confirmed using methods such as western blots using p-ser / p-thr specific antibodies, staining such as the Pro-Q Diamond stain used in this chapter or mass spectrometry used either to detect proportions of phosphorylated peptide fragments vs. un-phosphorylated fragments, or to analyse specific spots from 2D gels.

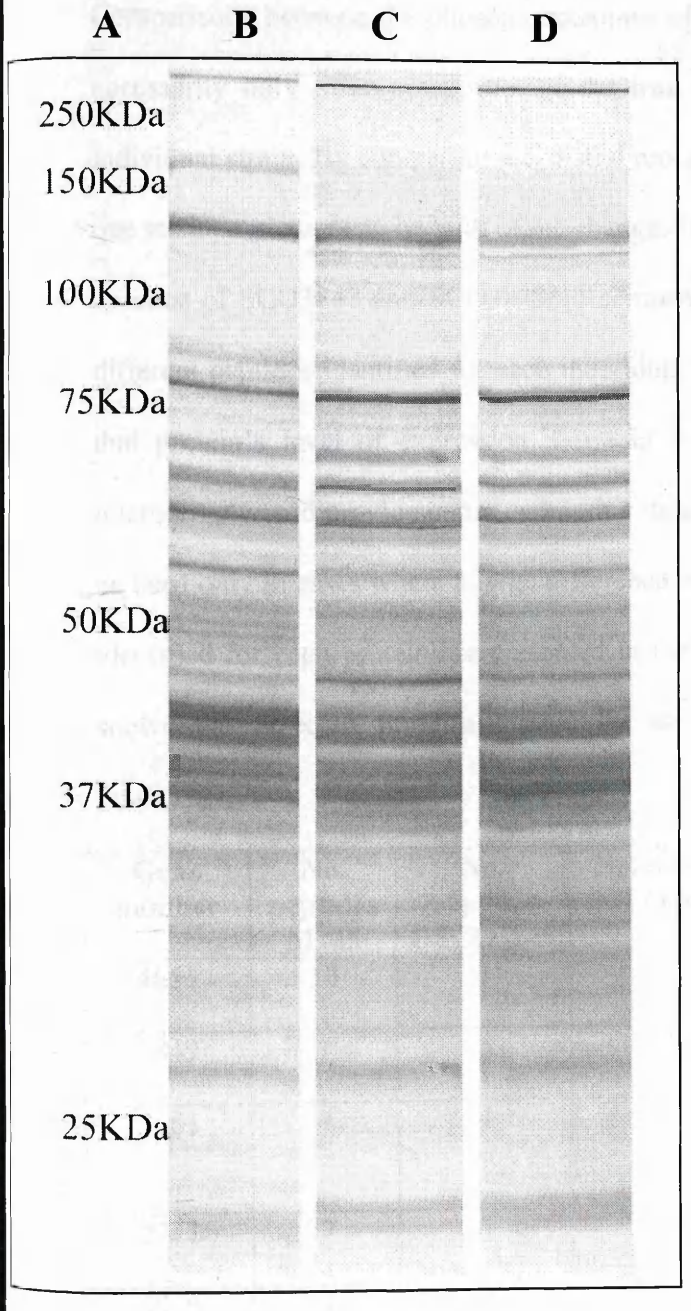


Figure 8.10: Pre-fractionating gel. A = Molecular size marker, B = phosphorylated proteins from parental M145 strain, C = phosphorylated proteins from DSCO3848, D = phosphorylated proteins from DSCO3843/44. All strains were grown on top of sterile cellophane membranes placed on NMMP media containing 10mM N – acetyl glucosamine for 48h at 30°C, this being the time point at which mutant phenotypes were most apparent. This image shows the three lanes placed next to each other, on the actual gel lanes were left empty between each sample to prevent cross – contamination.

8.7.4 Phospho proteome comparison between parent and mutants

Comparisons between the phosphoproteomes of the mutant and parental strains can necessarily only offer a snapshot of the true variation in the proteomes of each individual strain. By comparing the shared proteins to those that are found unique to one strain, a picture can be built of the changes in the phosphoproteome caused by the absence of SCO3848 and SCO3843/44 compared to the wild type. The numbers of different peptides identified for each individual protein can be taken as a measure of that protein's level of expression, however it is not the best method for this (a microarray would provide better expression data for example), therefore this data will be used only in cases where a large difference in expression is detected. The peptides identified for each protein are presented in the appendix for this chapter. Table 8.1 shows the detected proteins shared by all 3 strains (M145, DSCO3848 and DSCO3843/44).

Gene number	No. peptides M145	No. peptides DSCO3848	No. peptides DSCO3843/44	Function
4655	**10**	3	5	rpoC, DNA-directed RNA polymerase beta' chain
1849	9	5	**20**	cobN, cobalamin biosynthesis protein.
3661	4	4	2	clpB, ATP-dependent protease ATP-binding subunit
2777	7	4	6	accC, acetyl/propionyl CoA carboxylase alpha subunit
3970	4	2	6	Xaa-Pro aminopeptidase
5477	2	5	4	putative oligopeptide-binding lipoprotein
6451	10	9	7	nikA, putative nickel binding protein (Ahn et al. 2006)
1850	4	6	2	putative chelatase
4662	4	5	2	tufl, elongation factor TU-1

Table 8.1: Proteins common to parental and mutant strains detected by mass spectrometry, plus number of peptides detected.

The proteins detected that are common to all strains are annotated as being involved in various metabolic pathways related to housekeeping activities such as protein synthesis, and primary metabolism. It is interesting to note the greater abundance of RpoC in M145 than in the mutants, and also the greater abundance of CobN in DSCO3843/44 than in M145 or DSCO3848. The following tables display phosphorylated proteins common to 2 of the 3 strains, and their function as annotated by SCODB.

Gene number	No. peptides DSCO3848	No. peptides DSCO3843/44	Function
4729	2	4	rpoA, DNA-directed RNA polymerase alpha chain
4930	4	3	putative enoyl-CoA hydratase
1081	2	2	putative electron transfer flavoprotein, alpha subunit
3617	2	2	hypothetical protein
6042	2	2	conserved hypothetical protein
2637	2	4	Putative serine protease (putative secreted protein)
4296	2	2	groEL2, 60 kD chaperonin 2
5044	2	2	fumB, fumarate hydratase class I

Table 8.2: Proteins shared between DSCO3848 and DSCO3843/44 detected by mass spectrometry.

Gene Number	No. peptides M145	No. peptides DSO3848	Function
5737	2	2	gpsI, guanosine pentaphosphate synthetase/polyribonucleotide nucleotidyltransferase
2523	2	3	Putative ATPase

Table 8.3: Phosphorylated proteins detected common to M145 and DSCO3848 but not DSCO3843/44.

Gene Number	No. peptides M145	No. peptides DSCO3843/44	Function
2113	2	2	bfr, probable bacterioferritin
4654	10	6	rpoB, DNA-directed RNA polymerase beta chain
2949	2	2	MurA UDP-N-acetylglucosamine transferase
5706	2	2	Probable translational initiation factor
1662	1	2	Conserved hypothetical protein

Table 8.4: Phosphorylated proteins detected common to M145 and DSCO3843/44 but not DSCO3848. SCO2999 is not included, due to this portion of the DSCO3848 gel being missing from the final LC-MS/MS dataset.

The above tables show that where two strains share a phosphorylated protein species, the amount detected by LC-MS/MS is virtually the same. The functions of these proteins are fairly disparate, ranging from translational machinery to a part of an electron transfer system. Whether these proteins are direct targets for phosphorylation by the STPK or FHA domains remains to be discovered, however an idea of the change in regulation of proteins can be gleaned by study of the phosphorylated proteins found uniquely in each strain.

Gene Number	Function
2447	Hypothetical protein
2494	Probable pyruvate phosphate dikinase
1648	arc AAA ATPase
2776	Acetyl/propionyl CoA carboxylase, beta subunit
7516	htpG, heat shock protein
1296	Conserved hypothetical protein
0560	cpeB, catalase/peroxidase
5249	Putative nucleotide-binding protein
5373	atpD, ATP synthase beta chain
3122	Putative nucleotidyltransferase
1626	Putative cytochrome P450
5199	Conserved hypothetical protein
2599	Putative ribonuclease

Table 8.5: Gene numbers of phosphorylated proteins found exclusively in M145 and their annotated functions.

Gene Number	Function
6198	Putative secreted protein
3487	Putative hydrolase
5113	BldKB, putative ABC transport system lipoprotein
0379	katA, catalase
5838	Putative protease
4647	nusG, transcription antitermination protein
4039	Hypothetical protein
1662	Conserved hypothetical protein
1793	Putative Spo0M protein
4723	adk, adenylate kinase
1968	Putative secreted hydrolase

Table 8.6: Gene numbers of phosphorylated proteins found exclusively in DSCO3848 and their annotated functions

Gene Number	Function
2447	Hypothetical protein
6198	Putative secreted protein
3127	Phosphoenolpyruvate carboxylase
5281	Putative 2-oxoglutarate dehydrogenase
5706	Probable translational initiation factor
2620	Putative cell division trigger factor
6279	Putative diaminobutyrate-pyruvate aminotransferase
5838	Putative protease
3122	Putative nucleotidyltransferase
3928	thiC, probable thiamine biosynthesis protein
1230	Putative secreted tripeptidylaminopeptidase
1476	metK, S-adenosylmethionine synthetase
2736	citA, citrate synthase
1968	Putative secreted hydrolase
6199	Secreted esterase
1665	Conserved hypothetical protein
0958	Putative methionine synthase
1643	pcrA 20S proteasome alpha-subunit

Table 8.7: Gene numbers of phosphorylated proteins found exclusively in DCO3843/44 and their annotated functions.

Tables 8.4 – 8.7 present the unique complement of potentially phosphorylated proteins detected by LC-MS/MS for M145 and the two mutant strains. The majority of these proteins found in M145 appear to be involved in primary metabolism and protein synthesis and it is also noteworthy that few of these proteins have been investigated before other than via sequence comparison. DSCO3848 displays a complement of unique phosphorylated proteins that appear to have similarity of function to those discovered in M145. The presence of a putative sporulation factor, Spo0M is perhaps indicative of a change in the regulation of cell division proteins. The lack of FHA domains appears to result in an increased number of phosphorylated proteins compared to the two other strains in this study, and the specific proteins unique to this strain are again for the most part not fully annotated other than through

sequence comparison. The most significant protein unique to the FHA domain mutant is CitA, citrate synthase. This ma y have some significance, especially when the phenotype of DSCO3843/44 when grown on media containing citrate is taken into account (chapter 6). The preponderance of phosphorylated proteins dedicated to primary metabolism found here is in good agreement with the phosphoproteome of the related organism *Corynebacterium glutamicum* (Bendt et al. 2003). This study also showed phosphorylation of CitA.

Potentially, CitA contains 50 sites which could be used to modify the behaviour of that protein via phosphorylation and dephosphorylation of ser/thr residues. These residues are identified in table 8.8, as well as the 3 residues either side, as these have been identified as being integral to FHA domain binding specificity (Durocher et al. 2000).

Residue number (numbered from N-C)	Potential amino acid phosphorylated (S or T)	Adjacent residues
2	S	--M, DNS
5	S	SDN, VVL
16	T	GEY, YPV
22	S	VID, TVG
23	**T**	**IDS, VGD**
38	T	RAQ, GLV
42	T	GLV, LDS
45	S	TLD, GYG
50	T	YGN, AAY
55	S	AYK, AIT
58	T	SAI, YLD
81	S	AER, SFV
82	S	ERS, FVE
97	**T**	**ELP, VDE**
102	S	DEL, AFK
136	S	AML, SVV
137	S	MLS, VVS
140	S	SVV, ALS
143	S	SAL, TFY
144	T	ALS, FYQ
149	S	YQD, HNP
161	S	RNL, TIR
162	T	NLS, IRL
171	T	KLP, IAA
180	S	YKK, IGH
202	T	LRM, FSV
204	S	MTF, VPA
215	**T**	**LDP, VVA**
235	S	QNC, TST
236	T	NCS, STV
237	S	CST, TVR
238	**T**	**STS, VRL**
244	S	LVG, SQA
245	S	VGS, QAN
252	S	MFA, ISA
254	S	ASI, AGI
271	S	ANQ, VLE
288	S	DVD, FIR
329	S	DVL, ALG
334	S	LGK, DEL
350	S	HAL, DDY
356	S	YFV, RSL
358	S	VSR, LYP
367	T	DFY, GLI

378	T	GFP, EMF
382	**T**	**EMF, VLF**
406	S	EPG, RIG
416	T	QIY, GVV

Table 8.8: Ser/thr residues potentially modified by phosphorylation in CitA (coded by SCO2736). Also included are the adjacent 6 residues to each serine or threonine. Those highlighted in bold seem likely to interact with an FHA domain according to Durocher et al. (2000).

8.8 *Summary*

The biochemical properties of *S. coelicolor* PknB when overexpressed and purified from *E. coli* mimic those shown by its mycobacterial orthologues. This implies that the mode of action of these proteins will be similar. As with the mycobacterial orthologues, the kinase domain of *S. coelicolor* PknB autophosphorylates. The conditions required for its purification (highly denaturing) did not prevent a portion of the purified protein migrating in SDS PAGE gels at a much higher molecular weight than was expected. The same phenomenon was observed for the kinase domain expressed without PASTA domains. This could be due to the degree of phosphorylation of the protein. Examination of the total phosphoproteomes of the parental strain and DSCO3848 and DSCO3843/44 reveals that the mutant strains early progression to secondary metabolism is due to differences in the regulation of primary metabolism proteins and the cell's protein synthesis machinery. This provides further evidence for the involvement of SCO3843, SCO3844 and SCO3848 in a global regulatory network.

8.9 *References for Chapter 8*

Ahn, B.-E., J. Cha, E.-J. Lee, A.-R. Han, C. J. Thompson, and J.-H. Roe. 2006. Nur, a nickel-responsive regulator of the Fur family, regulates superoxide dismutases and nickel transport in Streptomyces coelicolor
doi:10.1111/j.1365-2958.2006.05065.x. Molecular Microbiology **59**:1848-1858.

Av-Gay, Y., S. Jamil, and S. J. Drews. 1999. Expression and characterization of the Mycobacterium tuberculosis serine/threonine protein kinase PknB. Infect Immun **67**:5676-5682.

Bendt, A. K., A. Burkovski, S. Schaffer, M. Bott, M. Farwick, and T. Hermann. 2003. Towards a phosphoproteome map of Corynebacterium glutamicum. PROTEOMICS **3**:1637-1646.

Boitel, B., M. Ortiz-Lombardia, R. Duran, F. Pompeo, S. T. Cole, C. Cervenansky, and P. M. Alzari. 2003. PknB kinase activity is regulated by phosphorylation in two Thr residues and dephosphorylation by PstP, the cognate phospho-Ser/Thr phosphatase, in Mycobacterium tuberculosis. Mol Microbiol **49**:1493-1508.

Carrio, M. M., and A. Villaverde. 2002. Construction and deconstruction of bacterial inclusion bodies. J Biotechnol **96**:3-12.

Chopra, P., B. Singh, R. Singh, R. Vohra, A. Koul, L. S. Meena, H. Koduri, M. Ghildiyal, P. Deol, T. K. Das, A. K. Tyagi, and Y. Singh. 2003. Phosphoprotein phosphatase of Mycobacterium tuberculosis dephosphorylates serine-threonine kinases PknA and PknB. Biochem Biophys Res Commun **311**:112-120.

Cowley, S., M. Ko, N. Pick, R. Chow, K. J. Downing, B. G. Gordhan, J. C. Betts, V. Mizrahi, D. A. Smith, R. W. Stokes, and Y. Av-Gay. 2004. The Mycobacterium tuberculosis protein serine/threonine kinase PknG is linked to cellular glutamate/glutamine levels and is important for growth in vivo. Molecular Microbiology **52**:1691-1702.

Dasgupta, A., P. Datta, M. Kundu, and J. Basu. 2006. The serine/threonine kinase PknB of Mycobacterium tuberculosis phosphorylates PBPA, a penicillin-binding protein required for cell division
10.1099/mic.0.28630-0. Microbiology **152**:493-504.

Durocher, D., I. A. Taylor, D. Sarbassova, L. F. Haire, S. L. Westcott, S. P. Jackson, S. J. Smerdon, and M. B. Yaffe. 2000. The molecular basis of FHA domain:phosphopeptide binding specificity and implications for phospho-dependent signaling mechanisms. Mol Cell **6**:1169-1182.

Garry L. Corthals, V. C. W., Denis F. Hochstrasser, Jean-Charles Sanchez,. 2000. The dynamic range of protein expression: A challenge for proteomic research. Electrophoresis **21**:1104-1115.

Gygi, S. P., G. L. Corthals, Y. Zhang, Y. Rochon, and R. Aebersold. 2000. Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology 10.1073/pnas.160270797. PNAS **97**:9390-9395.

Jin, H., and V. Pancholi. 2006. Identification and Biochemical Characterization of a Eukaryotic-type Serine/Threonine Kinase and its Cognate Phosphatase in Streptococcus pyogenes: Their Biological Functions and Substrate Identification. Journal of Molecular Biology **357**:1351-1372.

Macek, B., I. Mijakovic, J. V. Olsen, F. Gnad, C. Kumar, P. R. Jensen, and M. Mann. 2007. The Serine/Threonine/Tyrosine Phosphoproteome of the Model Bacterium Bacillus subtilis 10.1074/mcp.M600464-MCP200. Mol Cell Proteomics **6**:697-707.

Ortiz-Lombardia, M., F. Pompeo, B. Boitel, and P. M. Alzari. 2003. Crystal structure of the catalytic domain of the PknB serine/threonine kinase from Mycobacterium tuberculosis. J Biol Chem **278**:13094-13100.

Tusnady, G. E., and I. Simon. 2001. The HMMTOP transmembrane topology prediction server. Bioinformatics **17**:849-850.

Villaverde, A., and M. M. Carrio. 2003. Protein aggregation in recombinant bacteria: biological role of inclusion bodies. Biotechnol Lett **25**:1385-1395.

Villaverde, A., and M. Mar CarriÃ³. 2003. Protein aggregation in recombinant bacteria: biological role of inclusion bodies. Biotechnology Letters **25**:1385-1395.

Wehenkel, A., P. Fernandez, M. Bellinzoni, V. Catherinot, N. Barilone, G. Labesse, M. Jackson, and P. M. Alzari. The structure of PknB in complex with mitoxantrone, an ATP-competitive inhibitor, suggests a mode of protein kinase regulation in mycobacteria. FEBS Letters **In Press, Corrected Proof**.

Young, T. A., B. Delagoutte, J. A. Endrizzi, A. M. Falick, and T. Alber. 2003. Structure of Mycobacterium tuberculosis PknB supports a universal activation mechanism for Ser/Thr protein kinases. Nat Struct Biol **10**:168-174.

Appendix to Chapter 8

Confirmation of band identity:

Expressed total PknB (Peptides detected in bold type)

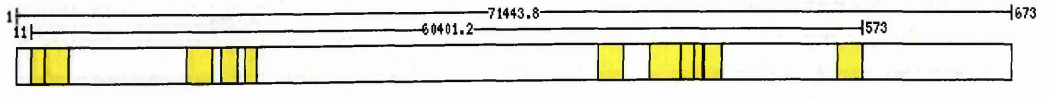
E. Coli proteins : chaperone Hsp70
succinate dehydrogenase

Streptomyces Coelicolor proteins :

>**gi|7481397**|pir||T36717 probable serine/threonine protein kinase - Streptomyces coelicolor gi|5102800|emb|CAB45215.1| (AL079308) putative serine/threonine protein kinase [Streptomyces coelicolor A3(2)]

MEEPRRLGGR **YELGPVLGRG GMAEVYHAHD TRLGR**QVAVK TLRADLARDP SFQARFRREA
QSAASLNHPA IVAVYDTGED
YIDNVSIPYI VMEYVDGSTL RELLHSGRKL LPER**TLEMTI GILQALEYSH R**AGIVHR**DIK**
PANVMLTRNG QVK**VMDFGIA**
RAMGDSGMTM TQTAAVIGTA QYLSPEQAKG EQVDARSDLY STGCLLYELL TVRPPFVGDS
PVAVAYQHVR EEPQAPSVFD
PEITPEMDAI VLKALVKDPD YRYQSADEMR VDIEACLDGQ PVGATAAMGA MAAGGYGAYP
DDQPTTALRS DGGGGATTML
PPMNPDDGGY GYDERPDRRR QQPRKKNTST IFLVLAGVLV LVGAILIGKY AFSGDGGPGN
DKVPVPAFIG LSK**ADAQQQA**
DNIDLVLTFK QQECEDQPKG NICAQDPK**QG TDVDKESTVN LVVSTGAPKV AVPNVIDKNI**
DEAKKQLEDK GFEVETKQTE
SSQDEGTILS QNPDPGKELE KGSTVTLEVA KAEEKATVPD VVGRTCDEAK AQVESGGDLT
AVCTDQPTND PNQVGK**VIST**
TPQSSTQVDP GSKVTIVVGK AVEKTKVPEV RGKTLAEARQ ILQQSGFTNV QVAQGSPGDD
NAKVFASNPQ PGSEVDDPAA
TPITLMTVPG DGGNGNGGNG NGGAIAGLPG FGD

Mass (mono): 71443.8 **Identifier:** gi|7481397 **Database:** D:/Xcalibur/database/StrepGREG.fasta
Protein Coverage: 144/673 = **21.4%** by amino acid count, 15685.1/71443.8 = **22.0%** by mass



Expressed STPK domain only

E. Coli proteins : chaperone Hsp70
integral membrane peptidase
succinate dehydrogenase
fumarase A = fumarate hydratase Class I
heat shock protein C 62.5

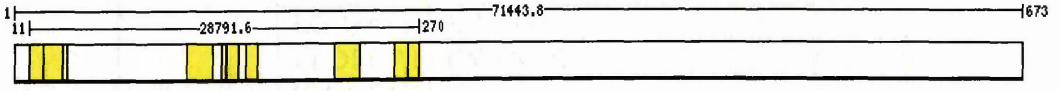
Streptomyces Coelicolor proteins :

gi|7481397|pir||T36717 probable serine/threonine protein kinase - Streptomyces coelicolor gi|5102800|emb|CAB45215.1| (AL079308) putative serine/threonine protein kinase [Streptomyces coelicolor A3(2)]

Peptides ID:

MEEPRRLGGR **YELGPVLGRG GMAEVYHAHD TRLGR**QVAVK TLRADLARDP SFQARFRREA QSAASLNHPA IVAVYDTGED
YIDNVSIPYI VMEYVDGSTL RELLHSGRKL LPER**TLEMTI GILQALEYSH R**AGIVHR**DIK PANVMLTR**NG QVK**VMDFGIA**
RAMGDSGMTM TQTAAVIGTA QYLSPEQAKG EQVDARSDLY STGCLLYELL TVR**PPFVGDS PVAVAYQHVR** EEPQAPSVFD
PEITPEMDAI VLK**ALVKDPD YRYQSADEMR** VDIEACLDGQ PVGATAAMGA MAAGGYGAYP DDQPTTALRS DGGGGATTML
PPMNPDDGGY GYDERPDRRR QQPRKKNTST IFLVLAGVLV LVGAILIGKY AFSGDGGPGN DKVPVPAFIG LSKADAQQQA
DNIDLVLTFK QQECEDQPKG NICAQDPKQG TDVDKESTVN LVVSTGAPKV AVPNVIDKNI DEAKKQLEDK GFEVETKQTE
SSQDEGTILS QNPDPGKELE KGSTVTLEVA KAEEKATVPD VVGRTCDEAK AQVESGGDLT AVCTDQPTND PNQVGKVIST
TPQSSTQVDP GSKVTIVVGK AVEKTKVPEV RGKTLAEARQ ILQQSGFTNV QVAQGSPGDD NAKVFASNPQ PGSEVDDPAA
TPITLMTVPG DGGNGNGGNG NGGAIAGLPG FGD

Mass (mono): 71443.8 **Identifier:** gi|7481397 **Database:** D:/Xcalibur/database/StrepGREG.fasta
Protein Coverage: 95/673 = **14.1%** by amino acid count, 10695.4/71443.8 = **15.0%** by mass



M145 Phosphoproteins:

SCO Number	Molecular weight (KDa)	Peptides Identified	Conserved Domains
2999	183.7	IADQLSTEATPEDLPAR	NAD-glutamate dehydrogenase
		IEFALQGGR	
		NNYAQNTAIANALAQSR	
		VQVVGEGGNLGLTQLGR	
		LLINLFEAR	

		APSGGADYLGDDAR	
		EQFAEQVDGHPLR	
		LGVEVTDERPYELR	
		VQDAFAATWTGK	
		LGVTQLMDR	
		HIFIDPTPDAATSYAER	
		MREETGASLEEIVR	
		GIHVVVHPQIVVR	
		ALEFLPTDR	
		LAENRPQGTAN	
		ATVHRPSYLDYVGVK	
		FDPQAIPDLPAPR	
		FLGLFSSAAYTESVR	
		DLLQILETYPR	
		EETGASLEEIVR	
		QAGSTFSQDYMEDTLR	
2113	19.2	VGQSVTEMFQADR	Probable bacterioferritin
		ILLLDGLPNYQR	
2447	16.0	VVVEGVEDEDGAR	Hypothetical protein
		LNWSPEPSPTR	
		TFEDSLDGVR	
4655	144.6	TADSGYLTR	DNA-directed RNA polymerase beta' chain
		AVQVHLVGEVQK	
		FPVGTIPPR	
		LESFDLDEEAER	
		TFHTGGVAGDDITQGLPR	
		FATSDLNDLYR	
		VVELFEAR	
		LGIQAFEPQLVEGK	
		LLDLGAPEIIVNNEK	
		IGLATADDIR	
1849	130.9	ALDGGYVPAGPSGSPTR	Cobalamin biosynthesis protein.
		AVAELDEPASSNFVK	
		LSAEYVFSPENR	
		GLWAEPDAETLER	
		YLVEGGPDNLTQLAR	
		DAVAGRPTVGVLFYR	
		FLSDTVLLTGEGFDEPR	
		VGNAVGLDTPASAVR	
		DAFPHVVGLIDDAVR	

4654	128.5	LHAGETETVGESGR	rpoB, DNA-directed RNA polymerase beta chain
		NASTANTNNAASTAPLR	
		EIVVETDDIDHFGNRR	
		GWDVSGLAEEWAQR	
		INPFGFIETPYR	
		ALMGANMMR	
		SPGVYFDSSIDK	
		EDELAGLLQHTIPNR	
		DIPNVSEEVLADLDER	
		EAAQTLLENLYFNPK	
3661	94.4	MVGATTLDEYRER	Probable clp-family ATP-binding subunit
		IVAGLPSVTGSTVAPPQPSR	
		MEIDSSPVEIDELQR	
		LDDLVVFSALSQEELSR	
2494	99.0	LILADTEAVR	Probable pyruvate phosphate dikinase
		MHAGAEDADELVEAVHR	
5737	79.4	ATINGIANPTMPEVGER	gpsI, guanosine pentaphosphate synthetase/polyribonucleotide nucleotidyltransferase
		VENVEDVLGVGAK	
2777	68.2	LREPEGDGVR	accC, acetyl/propionyl CoA carboxylase alpha subunit
		VTDPVEHTPR	
		VTVTLDGVR	
		VAAGEPLGFGQEDVR	
		IGPAPATESYLSVER	
		ASFGDDTLLVER	
		VIEEAPSVLLDEATR	
1648	65.2	DQIVALKEEVDR	arc AAA ATPase
		IVELQTNLAGVSAQNER	
		GLTIRPGDALLLEPR	
2776	57.4	VSGVTDHLAEDDPHALR	Acetyl/propionyl CoA carboxylase, beta subunit
		ANEEAHAALVQELR	
		LWDDGVIEPADTR	
7516	71.6	DNGIGMSYDEVTR	htpG, heat shock protein
		GVVDAQDLSLNVSR	
		YSDFITWPVR	
3970	54.0	SNDTEYSFR	Xaa-Pro aminopeptidase

		HSLAEAGELYGIPASDVR	
		IEDDILVTADGNR	
		YIEGTFFLR	
5477	65.3	ANLDNPVSGYIR	Putative oligopeptide-binding lipoprotein
		VWAQDVLSGGPTYLK	
1296	62.5	LMDELGYLTDAEQR	Conserved hypothetical protein
		LPDFSPAGQEALAELAR	
0560	80.8	VLGAGHQQSQLGVFTR	cpeB, catalase/peroxidase
		ADLVFGSNAELR	
6451	58.4	TVLDEKTNNTAR	Putative substrate binding protein
		AAAYGEIQK	
		RALDAAVDR	
		QDHALAYASDAK	
		SELDAVENVR	
		AVTLPSAGQVTGDR	
		ILDEAGWKPGSGGIR	
		VEGATWEVIEPR	
		FSDGEPLTAGDVVYTYR	
		PAYGPLPVDDPWYER	
5249	52.2	TEQTRPQQSLGTAAAR	Putative nucleotide-binding protein
		TGEQDQGVIGLR	
		ATGIPDEIEPSLSVR	
		SYEDDEVLSELAQR	
5373	52.2	VANNHDGVSVFAGVGER	atpD, ATP synthase beta chain
		GIYPAVDPLDSTSR	
		FTQAGSEVSTLLGR	
		TTTVETATATGR	
		TISMQPTDGLIR	
3122	49.8	TAVQEQQNGTGHAVR	Putative nucleotidyltransferase
		VGASVAGDHREIAGINNR	
1626	55.2	RADPADDIAGR	Putative cytochrome P450
		AGDGLLLGIAPGNVDPR	
1850	70.2	LAYDDDPAGFAAR	Putative chelatase
		TATALAAWAGR	
		LVGALDIER	
		QAALLALPHR	

2949	48.2	ADGQYLEAPQR	UDP-N-acetylglucosamine transferase
		VTGVDELGGYTHR	
5199	51.8	SHLFGAVDGYAR	Conserved hypothetical protein
		LWASLTDAHGVDGR	
4662	43.8	VNETVDIIGIK	tuf1, elongation factor TU-1
		GITISIAHVEYQTEAR	
		LLDEGQAGENVGLLLR	
		QVGVPYIVVALNK	
5706	105.7	ANPVQDAQGISIESR	Probable translational initiation factor
		IGAYQVATEVNDEER	
2599	146.1	TAAEGASEDELR	Putative ribonuclease
		LTSQVSLPGR	
1662	35.7	DGAGPGGPGGPGGHATADGGDGPAG	Conserved hypothetical protein
2523	33.6	SVTSANLAYQR	Putative ATPase
		SYAVDMVLDATAHPR	

DSCO3848 Phosphoproteins

SCO Number	Molecular weight	Peptides identified	Function
4655	145.0	IGLATADDIR	rpoC, DNA-directed RNA polymerase beta' chain
		VVELFEAR	
		DLPSLEAHVSVER	
1849	131.2	ALNGGYVPAGPSGSPTR	Cobalamin biosynthesis protein.
		VGNAVGLDTPASAVR	
		GLVNVLPTGR	
		YLSDTGEYPK	
		AVAELDEPASSNFVK	
5737	79.4	MEQQLDTLSPVTR	gpsI, guanosine pentaphosphate synthetase/polyribonucleotide nucleotidyltransferase
		GETQILGVTTLNMLR	
3661	94.4	LIGQTQAVR	ATP-dependent protease ATP-binding subunit
		IVAGLPSVTGSTVAPPQPSR	
		MVGATTLDEYR	
		IVAGLPSVTGSTVAPPQPSR	
6198	116.7	AAAAEADAQAAR	Putative secreted protein
		AQSLANTAASAAAAAR	
3487	88.6	TPEVVEAASK	Putative hydrolase
		YGGWLNGPR	
5477	65.3	FASTQDADSWDTTR	Putative oligopeptide-binding lipoprotein
		FFTNANDMDAR	
		KAVLYGADHVSLQTAR	
		AVLYGADHVSLQTAR	
		ANLDNPVSGYIR	
5044	60.7	GQNVLTEGGDEEALSR	fumB, fumarate hydratase class I
		LSLTGPLVVAR	
5113	65.5	IQQITDTAEATK	BldKB, putative ABC transport system lipoprotein
		YSSDSSYVDVTR	
2777	73.0	AAMGEAAVQAAR	accC, acetyl/propionyl CoA carboxylase alpha subunit
		SVAVFSDADADAR	
		VTDPVEHTPR	

		IGPAPATESYLSVER	
6451	58.4	TVLDEKTNNTAR	Putative substrate binding protein
		QDHALAYASDAK	
		SELDAVENVR	
		YANPAVDEALDAGR	
		ILDEAGWKPGSGGIR	
		TVLPIVPEHVAGK	
		VEGATWEVIEPR	
		FSDGEPLTAGDVVYTYR	
		PAYGPLPVDDPWYER	
3970	54.0	TLPISGTYSELQK	Xaa-Pro aminopeptidase
		HSLAEAGELYGIPASDVR	
0379	55.1	VDGNQGATPGVEPNSYGR	katA, catalase
		QMSPEQQQVLFENTAR	
5838	48.9	APAFADGEVER	Putative protease
		QGTEETVETIDSAAVR	
4662	43.5	LLDEGQAGENVGLLLR	tuf1, elongation factor TU-1
		QVGVPYIVVALNK	
		AVDEAIPEPER	
		ELLSEYEFPGDDVPVVK	
		GITISIAHVEYQTEAR	
2637	114.3	VQLVNTHGTVAGLGSVR	Putative serine protease (putative secreted protein)
		AYADPQPGVWEVEVEAR	
1850	70.2	TEVLAEDVR	Putative chelatase
		TATALAAWAGR	
		MVELPVGASEDR	
		LVGALDIER	
		LAYDDDPAGFAAR	
		TATALAAWAGR	
4647	32.8	ETPVELSFDQIQKN	nusG, transcription antitermination protein
4729	36.7	IPVDSIYSPVLK	rpoA, DNA-directed RNA polymerase alpha chain
		TLLSSIPGAAVTSIR	
4039	19.2	IYVSDAAAAAR	Hypothetical protein
		MYFAALLAR	

		SSYGEILLTDELLGR	
1662	35.7	DGAGPGGPGGPGGHATADGGDGPAG	Conserved hypothetical protein
		PVPISGTTSDADLAR	
		VPLGELPEDAR	
4930	28.1	DREEALELAAR	Putative enoyl-CoA hydratase
		AAELVFTAR	
		ELGLVDQLVPEGR	
		FGEFVLVR	
2523	33.6	SVTSANLAYQR	Putative ATPase
		GLHSYLEGDTSEPVR	
		SYAVDMVLDATAHPR	
1081	32.4	ESTGRPELTEAAIVVSGGR	Putative electron transfer flavoprotein, alpha subunit
		KPTLELLTLAR	
3617	24.3	ALQSVVAHVR	Hypothetical protein
		ISDDLTDVITDLR	
6042	28.1	SHIPAFLADETPR	Putative chlorite dismutase
		AGYADEVQELFDQLAAK	
1793	28.4	AVDSGDLDAINVHPLPAQK	Putative Spo0M protein
		GMNIGVTTELEIAR	
4296	56.8	AEIENSDSDYDR	groEL2, 60 kD chaperonin 2
		MEASLDDPYILIANSK	
4723	24.0	MEQPDAEGGFLLDGFPR	adk, adenylate kinase
		LAETLHIPHISTGDLFR	
1968	31.4	GASGYAPENTLAAVDR	Putative secreted hydrolase
		LVVQSFDADSIR	

DSCO3843/44 Phosphoproteins

SCO Number	Molecular weight (KDa)	Peptides identified	Function
2999	183.7	APSGGADYLGDDAR	Putative NAD-glutamate dehydrogenase
		IADQLSTEATPEDLPAR	
		NNYAQNTAIANALAQSR	
		LGVEVTDERPYELR	
		LGVTQLMDR	
		GIHVVVHPQIVVR	
		QSDAFDLSNLSVAMR	
		EQFAEQVDGHPLR	
		ALEFLPTDR	
		LLINLFEAR	
4655	145.0	IGLATADDIR	DNA-directed RNA polymerase beta' chain (fragment)
		LESFDLDEEAER	
		TFHTGGVAGDDITQGLPR	
		DLPSLEAHVSVER	
		VVELFEAR	
1849	131.0	VGNAVGLDTPASAVR	Cobalamin biosynthesis protein.
		GLVNVLPTGR	
		GLWAEPDAQTLER	
		YLVEGGPDNLTQLAR	
		FLSDTVLLTGEGFDEPR	
		FAEGMELR	
		AVAELDEPASSNFVK	
		LSAEYVFSPENR	
		VTGFEVVPVAELGR	
		LSWEVGQSLADSLIQR	
		EQGPDEVPVYVADPER	
		DGLHILGGGPVGEPR	
		YLVEGGPDNLTQLAR	
		DAVAGRPTVGVLFYR	
		FLSDTVLLTGEGFDEPR	
		VGNAVGLDTPASAVR	
		GHATVVDHLVPPMAR	
		ASQVWGGQANALPGLR	
		DAFPHVVGLIDDAVR	
		LSWEVGQSLADSLIQR	
2113	19.2	VGQSVTEMFQADR	bfr, probable bacterioferritin
		ILLLDGLPNYQR	

2447	16.0	VVVEGVEDEDGAR	Hypothetical protein
		LNWSPEPSPTR	
4654	128.5	LHAGETETVGESGR	rpoB, DNA-directed RNA polymerase beta chain
		STANTNNAASTAPLR	
		NASTANTNNAASTAPLR	
		VESALESGQDVPTK	
		EIVVETDDIDHFGNR	
		INPFGFIETPYR	
6198	116.7	AAAAEADAQAAR	Putative secreted protein
		TAVQASSAAHNAAR	
		EMAAAEAEDR	
		AEAAYASGDTASALANGR	
		AQSLANTAASAAAAAR	
3127	101.3	VVRPEPADEAR	Phosphoenolpyruvate carboxylase
		GDVQEVMLGYSDSSK	
		LGDLLGETLVR	
3661	93.1	IVAGLPSVTGSTVAPPQPSR	clpB ATP-dependent protease ATP-binding subunit
		MEIDSSPVEIDELQR	
6198	116.7	AAAAEADAQAAR	Putative secreted protein
		TAVQASSAAHNAAR	
5281	139.2	GLTDTAIIR	Putative 2-oxoglutarate dehydrogenase
		IAESQVNIPDHITVHPR	
2777	68.2	VTVTLDGVR	accC, acetyl/propionyl CoA carboxylase alpha subunit
		VAAGEPLGFGQEDVR	
		IGPAPATESYLSVER	
		LLAHPAVVAGELDTGLVER	
		AAMGEAAVQAAR	
		MFDTVLVANR	
5477	65.3	FFTNANDMDAR	Putative oligopeptide-binding lipoprotein
		AVLYGADHVSLQTAR	
		ANLDNPVSGYIR	
		VWAQDVLSGGPTYLK	
5706	105.7	ANPVQDAQGISIESR	Probable translational initiation factor
		IGAYQVATEVNDEER	

2620	51.2	QYDQATQAQER	Putative cell division trigger factor
		LLEDEINTR	
6451	58.4	TVLDEKTNNTAR	Putative substrate binding protein
		AAAYGEIQK	
		RALDAAVDR	
		SELDAVENVR	
		ILDEAGWKPGSGGIR	
		FSDGEPLTAGDVVYTYR	
		PAYGPLPVDDPWYER	
6279	56.9	RNEEQLAR	Putative diaminobutyrate-pyruvate aminotransferase
		VFPTSSAVNTLR	
		VAGSADELIADVR	
		IPVVVDEIQSGMGR	
		LSALTEEIDAHLEEVR	
3970	54.0	IEDDILVTADGNR	Xaa-Pro aminopeptidase
		LVEWGLVEGPVER	
		YIEGTFFLR	
		VLDRAEATSER	
		SNDTEYSFR	
		HSLAEAGELYGIPASDVR	
5838	48.9	APAFADGEVER	Putative protease
		QGTEETVETIDSAAVR	
		NRLDEIPHELANPSR	
5044	60.8	GQNVLTEGGDEEALSR	fumB, fumarate hydratase class I
		LVTAEGVATFEADGR	
3122	49.8	TAVQEQQNGTGHAVR	Putative nucleotidyltransferase
		VGASVAGDHREIAGINNR	
2949	48.2	ADGQYLEAPQR	murA, UDP-N-acetylglucosamine transferase
		VTGVDELGGYTHR	
3928	67.2	WGADTVMDLSTGR	thiC, probable thiamine biosynthesis protein
		NSPVPIGTVPLYQALEK	
1230	58.5	SGEMLPHMTTPNTAR	Putative secreted tripeptidylaminopeptidase
		QGALIYNPGGPGGSGLR	

1476	43.5	AYADIANLVR AIDEVFDLRPAAIIR	metK, S-adenosylmethionine synthetase
2637	114.3	NSTESYFVTVPEGAK VQLVNTHGTVAGLGSVR ALEVAIGGLR AYADPQPGVWEVEVEAR	Putative serine protease (putative secreted protein)
4662	43.5	GITISIAHVEYQTEAR LLDEGQAGENVGLLLR	tuf1, elongation factor TU-1
2736	47.7	AQTGLVTLDSGYGNTAAYK AMGFPTEMFTVLFALGR	citA, citrate synthase
4729	36.5	IPVDSIYSPVLK PSLTEEVVDEFR AADLALPIEELELTVR IDGVLHEFTTVPGVK	rpoA, DNA-directed RNA polymerase alpha chain
1850	70.2	LAYDDDPAGFAAR TATALAAWAGR	Putative chelatase
5706	105.7	ANPVQDAQGISIESR IGAYQVATEVNDEER	Probable translational initiation factor
1662	35.7	DGAGPGGPGGPGGHATADGGDGPAG VPLGELPEDAR	Conserved hypothetical protein
4296	56.8	AEIENSDSDYDREK MEASLDDPYILIANSK	groEL2, 60 kD chaperonin 2
1968	31.4	GASGYAPENTLAAVDR VPTLEQYVHR TRDGELVVLHDESLAR LVVQSFDADSIR DGELVVLHDESLAR LDAGSWFGSEYAGAR	Putative secreted hydrolase
4930	28.1	DREEALELAAR AAELVFTAR FGEFVLVR	Putative enoyl-CoA hydratase
3617	24.3	ALQSVVAHVR ISDDLTDVITDLR	Hypothetical protein

1081	32.3	ESTGRPELTEAAIVVSGGR KPTLELLTLAR	Putative electron transfer flavoprotein, alpha subunit
6042	28.1	SHIPAFLADETPR AGYADEVQELFDQLAAK	Putative chlorite dismutase
6199	74.4	ADLTGTGGAVR TVNGTQQLAALSSR	Secreted esterase
1665	21.4	FVAGTVGLPGR LLQDEENGPPMLR	Conserved hypothetical protein
0985	83.7	TPAELNALER LLPVYAEVLTDLR	Putative methionine synthase
1643	27.7	GYTYDRDDVTAR EIPAERLEVAVLDR	pcrA 20S proteasome alpha-subunit

9 Discussion

9.1 *Preliminary remarks*

It is the aim of this thesis to present a plausible role for the conserved signalling genes found within the *pknB* gene cluster. The preceding results Chapters have focussed on the evolutionary relationships between these genes, and also on the phenotypic and proteomic results of their absence. All these areas could have been explored in much greater depth, however the author believes that this would have been at the expense of the opportunity to gain a real insight into the overall function of a gene cluster that is becoming the focus of an increasing amount of clinical interest.

9.2 *Evolution of the pknB gene cluster*

9.2.1 Origin of the cluster

Chapter 3 of this thesis presents G+C content for all the genes present in this cluster, as well as phylogenetic trees for the conserved actinobacterial genes. This analysis showed a G+C content and evolutionary relationship between conserved genes that was mirrored in an accepted 'molecular clock' gene, that for the 16S subunit of RNA polymerase (Kurland, Canback, and Berg 2003). This adds to the evidence that these genes are of ancient origin, and when the distribution between actinobacteria and firmicutes is taken into account, it seems highly plausible that the cluster was present in the original ancestor of these clades. This evidence makes the possibility of horizontal transfer from eukaryotic sources increasingly unlikely. When compared to a genuine case of eukaryote – prokaryote horizontal gene transfer (HGT), α – amylases (Da Lage, Feller, and Janecek

2004), it becomes increasingly tenuous to hold that this was the case with the *pknB* cluster. The criteria of Da Lage et al state that the genes must be few in number and found in organisms that are evolutionarily distant from each other, however this is not the case with *pknB*, as this particular gene is commonly found exclusively in relatively close Gram – positive organisms. Subsequent evolution away from the common ancestor resulted in the different arrangements of genes shown in Chapter 3. It seems likely that at one point early on in the history of the firmicutes there existed an organism containing the 'complete' *pknB* cluster. This cluster was then brought together with the essential *priA* associated genes (as shown in *D. hafniense, D. reducens* and *M. thermoacetica*) before a further deletion of genes resulted in the conservation of the two versions of the cluster within the firmicutes. It seems plausible that the association of *pknB* with *priA* occurred at the same point in evolutionary history as the divergence of the actinobacteria and firmicutes, indicating that these two clades were formed after a dramatic genome rearrangement in the common ancestor. The duplication of *pknB* like genes in the actinobacteria (resulting in SCO2110 and orthologues) probably occurred just after the split between the actinobacteria and firmicutes, as the duplication of the STPK domain is carried in all sequenced actinobacteria, however there is no corresponding duplication in the sequenced firmicutes.

9.2.2 Why do the genes of the *pknB* cluster group together?

The above question has two answers, both equally plausible from an evolutionary perspective: an evolutionary advantage is gained from either their co – transcription, or advantage is gained through co – incidental clustering of essential (or near – essential) genes. As mentioned in Chapter 3, the distribution of this cluster extends to the

actinobacteria and firmicutes, with much greater difference in gene position found in the firmicutes. There is a great deal of evidence showing genes that cluster together are often functionally connected (Huynen et al. 2000; Zheng, Roberts, and Kasif 2002), and that gene order need not necessarily be conserved. This is the essence of the 'über – operon' theory expounded by Bork and co – workers (Lathe, Snel, and Bork 2000). The *pknB* gene cluster is a good candidate for being classified as an über – operon, as it is well conserved across disparate species, however the number of genes in the cluster does show some variation, as does the gene distribution in *T. whipplei*. Insufficient data is available regarding expression and regulation of these genes in all the organisms in which they appear, however the high degree of conservation of gene order allows a confident prediction of conservation of regulation and function (Zheng, Roberts, and Kasif 2002). The *pknB* cluster adds a further layer of complexity to the picture by containing genes that are involved with two different and potentially unrelated cellular processes (signalling and cell division). Current evidence for a functional link between the two processes rests on the discovery by Dasgupta et. al. that PknB in *M. tuberculosis* is able to phosphorylate the penicillin binding protein transcribed from the *pbpA* gene (orthologous to SCO3847) (Dasgupta et al. 2006). While this is encouraging evidence, it should be remembered that PknB is able to phosphorylate a wide variety of proteins, and it is yet to be proven that this relationship reflects the reality in actinobacteria. The possibility of two functionally unrelated groups of genes being co – selected over evolutionary time is feasible. This is shown most clearly in the firmicutes, where the *pknB* orthologue and the cognate phosphatase gene maintain a genomic relationship with *priA*, despite there being little in the way of experimental evidence which would suggest

they are functionally related. The clustering of functionally unrelated essential or near – essential genes is seen in the genomes of the *Streptomyces*, where the core section of the linear genomes contain the more essential genes, and the arms contain the less essential genes (Bentley et al. 2002; Ikeda et al. 2003). The selective pressure behind this arrangement is demonstrated by the fact that genome rearrangement is far more prevalent in the arms of the chromosome when compared to the core region (Chen et al. 2002). It is plausible that the same effect is observed within genomes at the level of the individual genes, where clustering of more important genes confers selective advantage by virtue of their mutual preservation during genome rearrangement.

Where does this leave the *pknB* cluster? The presence of the PASTA domains on PknB coupled to the presence of *rodA, pbpA* and *crgA* (Del Sol et al. 2003; Del Sol et al. 2006) strongly suggests an involvement in cell division. The probable co – transcription of *pstP*, *rodA,* and *pbpA* also suggests a functional relationship between serine – threonine phosphosignalling and cell division. The lack of *in vivo* evidence for this speculation forces alternative explanations for the existence of the cluster, the most likely of which is the mutual preservation theory outlined above. It is the author's opinion that there is an as – yet unproven functional linkage between these genes, and that the conservation of the cluster is due to a combination of the mutual preservation theory (especially in the firmicutes) and the tendency for functionally related genes to cluster together.

9.3 *Evolution of pknB: Functional implications*

9.3.1 Differential selective pressure either side of the membrane

The discovery that the STPK – coding domains of *pknB* have been under stronger purifying selection compared to the PASTA domains shows that it is possible for different functional domains to evolve at different rates on the same gene. This then raises the question of why this should be so.

The implication is that the function of orthologous STPK domains is conserved, and this is reflected in both the inter - genomic conservation and degree of purifying selection under which it is found. The surprising result for the PASTA domains is that they are under significantly more positive selective pressure than the STPK domains, implying that the PASTA domains have different functions than their orthologous domains. This is a conclusion which is at odds with the conserved structure of the PASTA domains, which implies a conserved mode of action. The challenge is to rationalise the conservation of structure and implied function with the very low level of sequence conservation between PASTA domains.

9.3.2 Position – dependent evolution of PASTA domains

The discovery that the PASTA domains are phylogenetically more related to their orthologues in different species than their immediate neighbours on the same gene is good evidence against the hypothesis that these domains are a series of repeats of a single domain. The implication of this position – dependent clustering is that each *pknB* and

pbp2X like gene codes for proteins that require specific numbers of PASTA domains of specific types. When taken in conjunction with the predicted peptidoglycan monomer binding function of the PASTA domains (Yeats, Finn, and Bateman 2002), the high K_a/K_s values imply that specific peptidoglycan monomers bind to specific PASTA domains. The most variable part of the peptidoglycan molecule is the stem peptide (Schleifer and Kandler 1972), so it seems likely that it is this variation that drives the positive selection of the PASTA domains. The theory is further supported when the *pbp2X* associated PASTA domains are taken into account. These too show position – dependent clustering on phylogenetic trees, and these domains do not cluster with any domains found on STPKs. Therefore, each genus of firmicutes and actinobacteria contains a unique complement of independent PASTA domains, likely to be suited to the genus – specific features of the Gram – positive cell wall. When the second *Streptomyces pknB* – like gene is taken into account, the phylogenetic relationship supports the above theory. Instead of clustering with PASTA domains within the same species on the 'second' gene (e.g. SCO3848 domain 1 clustering with SCO2110 domain 1), domains associated with orthologous domains in different species (e.g. SCO3848 domain 1 associates with SAV4338 domain 1). Thus the two proteins which share a PknB – like transmembrane configuration in *Streptomyces* species are likely to be sensing 8 different peptidoglycan moieties.

9.3.3 Linking evolution, structure and function of PASTA domains

The method used to look at the overall selection pressure in two different domains of a gene can be further focussed to examine the selection pressure on individual codons by use of sliding window analysis. This method allows a correlation between amino acid

conservation and the selection pressure found on their codons. In *S. pneumoniae*, the PBP2X protein has been studied in relation to penicillin resistance (Dessen et al. 2001; Granger et al. 2005). This work has led to the discovery that penicillin resistant isolates have mutations in their PASTA domains. This information naturally leads to the assumption that these particular amino acids interact with beta – lactam antibiotics and their structurally similar relations, the cell wall NAM and NAG based muropeptides. Little experimental evidence exists showing similar effects for orthologous PASTA domains in other members of the firmicutes, however comparison of aligned sequences in terms of selection pressure allows comparisons to be made. The results for PBP2X PASTA domain 1 shown in Chapter 4 show that there is a close correlation between areas of the corresponding genes under positive selection and areas of the protein which show amino acid substitutions in penicillin resistant isolates. Therefore, it is possible to conclude that these areas correspond to segments of the PASTA domain proven to interact with cell wall analogues, and their universal variability supports the theory of specificity for a unique pool of pepdidoglycan monomers.

9.3.4 Variability In Gram – positive peptidoglycan: A clue to PASTA domain diversity and function.

The preceding sections have shown that the PASTA domains are likely to sense pools of unlinked peptidoglycan monomers during a process linked to cell division, consistent with the conservation of *pknB* like genes in a gene cluster containing genes likely to have a cell division function. The great variability of PASTA domains would be difficult to explain if they were found in Gram – negative organisms, as the variety of peptidoglycan monomers is restricted to 3 types (Schleifer and Kandler 1972). It is instructive to

compare the PASTA domain containing PBPs of Gram – positive bacteria with the PBPs of Gram – negative bacteria, as localisation of cell division machinery does not use a pool of unlinked monomers. Indeed, this is the case for the majority of the Gram – positive cell division proteins as well. Why has the ability to localise the cell division machinery using the cell wall as a signalling device been used when adequate mechanisms exist which do not need this system?

The best explanation for a role for the PASTA domains comes from studies performed in *S. pneumoniae*. These studies showed PBP2x localising to the division septum after the FtsZ ring had been formed, and that this was associated with the formation of a second ring based on FtsW. The D,D-carboxypeptidase PBP3 was found to localise away from the division septum, and it was postulated that the localisation of PBP2x was guided by the availability of unlinked peptidoglycan units (i.e. where carboxypeptidase activity was minimal) (Morlot et al. 2003; Morlot et al. 2004). The PASTA domains are ideal candidates for guiding the localisation of PBP2x to the division septum. The variability in PASTA domain amino acid sequence putatively reflects the stem peptide multiplicity of unlinked monomers which could be present (Garcia-Bustos and Tomasz 1990), with specific regions of each domain interacting with specific moieties. This system offers an alternative, reliable method to localise the appropriate proteins to the site where they are required. This is advantageous in cases where one or more members of a complex septal ring system are not present to be bound to. The independence of this system would ensure that even damaged cells would still be able to achieve cell division, thus achieving a selective advantage.

The involvement of PBP2x in cell division is beyond doubt, however the STPKs have a very different biochemical function with no obvious direct role in the division / elongation of the bacterial cell. This is puzzling when the fact that these kinases possess multiple peptidoglycan – sensing PASTA domains, and are genetically linked to genes involved in the manufacture of peptidoglycan in the case of the actinobacteria and a subset of the firmicutes. Why should proteins predominantly associated with the phosphorylation of serine and threonine residues need the ability to localise to particular areas of cell wall activity?

9.4 *SCO3848*

9.4.1 Mutant phenotype: a clue to function

DSCO3848 has a phenotype defined by early aerial development and antibiotic production when grown on rich glucose – containing media such as NE. It should also be noted that the phenotype is not seen on SFM, which does not contain glucose. A potential red herring is the phenotype observed on NMMPG when compared to NMMP (in which DSCO3848 shows retarded aerial development compared to the parent) however this is likely to be due to a greater sensitivity to the repressive effect of the phosphate buffer used in NMMP. When this buffer is used in NE media, aerial development is repressed in DSCO3848 as well. The addition of more complex disaccharide carbon sources did not result in phenotypic differences compared to the parental M145 strain when the strains were grown on minimal media. Increasing the concentration of tryptone also caused an increase in antibiotic production. Commercial tryptone preparations contain a percentage of sodium chloride, typically 0.5 – 1.0%. To test whether the increase in salt levels was

causing the increase in secondary metabolism or not, the mutant and parental strain were compared whilst growing on 2xYT media containing varying salt concentrations. The results indicate that increasing salt concentrations has a repressive effect on secondary metabolism in both strains.

The *S. coelicolor* genome codes for 140 secreted cellulases, proteases, chitinases and hydrolases (Bentley et al. 2002). The purpose of these proteins is presumably to break down external complex molecules into more readily assimilated small carbon compounds that need minimal processing before being fed into various metabolic cycles. Tryptone contains a complex mixture of peptides that would presumably be broken down by the proteases before being transported into the cell. The increase in secondary metabolism seen by DSCO3848 in situations of increased tryptone concentration may therefore be assigned to the same mechanism that increases secondary metabolism for this strain in glucose supplemented media.

9.4.2 Is SCO3848 involved in primary or secondary metabolism?

The evidence so far shows that SCO3848 has a role in the regulation of secondary metabolism, however, is this role a direct regulatory one, or is it a case of lack of regulation of primary metabolism causing 'overflow' metabolism (Hodgson 2000) to take place? The analysis of the phosphoproteomes from DSCO3848 and M145 shows that no proteins which are directly involved in the regulation of secondary metabolism are differently phosphorylated in the two strains. Therefore, an explanation is needed for the obvious metabolic change seen in *S. coelicolor* when SCO3848 is disrupted. From studies

performed in the firmicutes and *mycobacteria* (see introduction), the majority of proteomic targets are either FHA domains, or they are proteins associated with the regulation of gene expression. No study has thus far shown the direct phosphorylation of any protein directly involved in metabolism. The dilemma here is to rationalise this lack of direct regulatory evidence with the obvious acceleration of the life cycle seen in DSCO3848. The most plausible speculation for the role of SCO3848 is that it regulates the timing of secondary metabolism activation at a regulatory level upstream to that of the individual metabolic enzymes.

9.4.3 Timing of SCO3848 transcription: A clue to function.

The transcription of SCO3848 coincides with the growth of the early aerial mycelium. This period of growth does not involve a large amount of septum formation and cell division. Therefore, either SCO3848 is involved in the formation of the few septa found in the hyphae at this point, or it is involved in the extension of hyphae. One gene with which it is clustered, *crgA* (SCO3853), has been shown to be involved in the extension of the hyphal tip through repression of the formation of the FtsZ ring (Del Sol et al. 2006). The expression of *crgA* is concurrent with that of SCO3848. Their genomic linkage and similar expression patterns aid the hypothesis that the expressed proteins are involved in the same process. The genes encoding RodA / FtsW and PBP proteins can be plausibly assigned to cell elongation processes due to their homology with proteins involved in this process in other bacteria (Boyle et al. 1997; Khattar et al. 1997; Pastoret et al. 2004). It is therefore postulated that SCO3848 uses the presence of unlinked peptidoglycan monomers specific to this period of hyphal growth as a signalling device to stimulate the regulation of a variety of cellular processes dedicated to the maintenance of primary

metabolism during the period of aerial hyphal extension. It is also interesting to note the coincidental transcription of both SCO3848 and the global regulator of secondary metabolism *ramR* (Nguyen et al. 2002). The RamR protein is a response regulator that is analogous to Spo0A in *B. subtilis* in that it appears to integrate various extracellular signals and then act as a regulator of secondary metabolism via control of production of a lantibiotic (Kodani et al. 2004) and transcription of a diverse set of associated genes (Paolo et al. 2006). The temptation is to assign SCO3848 a role in the RamR mediated control of secondary metabolism (presumably as one of the sensors of extracellular conditions) however without evidence this must remain speculation.

9.4.4 Redundancy

The attention of this discussion has so far focussed on SCO3848. There is another transmembrane STPK with four PASTA domains in the extracellular portion found in the *S. coelicolor* genome. This is SCO2110, and its genomic position has been described in Chapter 3 of this thesis. Unpublished results (Jung and Buttner, personal communication) from disruption of both SCO2110 and SCO3848 show an additive effect in terms of antibiotic production in *S. coelicolor* M600. This additive effect, coupled to the structural similarity of the proteins expressed from these two genes leads easily into the assumption that they are performing similar functions. This is, however, partially refuted by the evidence shown in Chapter 4 which shows that orthologues are more closely related than homologues in both STPK and PASTA domains. There is enough similarity between the two STPK proteins to postulate that function is similar, if not identical. The effect of the possession of redundant genes is seen in the disruption of the orthologue of SCO3848 in *S. avermitilis*, SAV4338. This disruption did not cause any observable phenotypic

difference to that of the wild type. This is likely to be due to redundancy within the genome, with SAV4662 (othologous to SCO2110) being the primary candidate for duplicating SAV4338 function due to its similar expressed protein structure.

9.4.5 Microscopic analysis: is lack of change important?

The lack of obvious phenotypic change in hyphae produced in DSCO3848 is at first disconcerting. After all, the loss of a gene so closely associated evolutionarily and genomically with an essential cell wall gene such as *rodA/ftsW* would indicate some kind of relationship. The answer must lie in the particular use to which PknB like proteins put their PASTA domains. If these domains are used to sense peptidoglycan monomers associated with a certain phase of cell wall development, is this a physical or temporal signal, in other words are the PASTA domains being used to localise the STPK to a specific point on the cell wall, or are they being used to ensure that a particular signalling pathway is activated during a specific phase of cell wall development? The answer could involve both ideas, however a localisation role for the PASTA domains would imply that other localised cell wall related proteins required modification on ser / thr residues. It seems likely that PknB like proteins would localise with the products of other genes with which it is clustered (Wolf et al. 2001; Rogozin et al. 2002), and these gene products have been predicted to be involved in cell elongation (Boyle et al. 1997; Khattar et al. 1997; Pastoret et al. 2004). As elongation of aerial hyphae occurs during the same time at which SCO3848 is expressed, it seems plausible that the two processes are linked. The lack of obvious phenotypic changes to the hyphae when viewed microscopically implies that the role of the PknB like STPKs is not in the regulation of the structure of the cell wall, it is more likely they are involved in maintaining intracellular conditions conducive

to the elongation process. The PASTA domains are likely to be used to sense the state of the cell wall and, once elongation is over, trigger a change in cell metabolism that is conducive to secondary metabolism and spore formation. Electron micrographs show that a SCO3848 mutant version of *S. coelicolor* M600 has infrequent swollen spores (Jung and Buttner, personal communication). This effect could be due to the premature formation of spores in hyphae still undergoing elongation, as the signal for an orderly switch from elongation to sporulation is disrupted by the absence of the PknB protein. It is interesting to compare the differences in morphology between *pknB* mutants in *Streptomyces* and in species of *Mycobacteria* with depleted amounts of PknB (this being an essential protein in *Mycobacteria* (Sassetti, Boyd, and Rubin 2003; Fernandez et al. 2006)). When PknB is depleted in *Mycobacteria*, the cells take on a thin, elongated phenotype that contrasts to the swollen phenotype observed when the protein is overexpressed (Kang et al. 2005). This is a very strong indicator that PknB is involved in the regulation of cell shape, and this was supported by evidence from the same study that showed PknB to phosphorylate the mycobacterial orthologue of DivIVA. DivIVA is an essential cell elongation specific protein in *S. coelicolor* (Flardh 2003), and so this implicates PknB in the elongatory phase of the cell cycle. However, the PknB phosphorylation site present in mycobacterial DivIVA is not present in the *S. coelicolor* protein, and so it seems that this particular signalling pathway differs between the *Streptomyces* and *Mycobacteria*.

9.4.6 Why is DSCO3848 more sensitive to phosphate than M145?

When grown on media buffered with phosphate based solutions such as NMMPG, the effect of phosphate repression of secondary metabolism (Doull and Vining 1990) is

exaggerated in DSCO3848 when compared to M145. There is no obvious link between phosphate availability and the state of peptidoglycan beyond the most general requirements for metabolism. There has been a proven link between a ser / thr signalling cascade and the control of secondary metabolism in the form of the AfsR cascade, in which a global regulator of actinorhodin production (AfsR) is phosphorylated by multiple STPKs (Sawai et al. 2004) before going on to facilitate the binding of RNA polymerase to the promoter of *afsS*. The AfsS protein then stimulates antibiotic production and secondary metabolism (Umeyama, Lee, and Horinouchi 2002). AfsS transcription is also directly influenced by the PhoR – PhoP system, whereby the presence of phosphorylated PhoP during times of depleted extracellular phosphate facilitates transcription of AfsS through binding of phosphorylated PhoP to the promoter region of *afsS* (Sola-Landa, Moura, and Martin 2003). As SCO3848 seems to facilitate the production of secondary metabolites in the presence of phosphate, it may be that a complementary STPK system balances the AfsS response to phosphate limitation, allowing other environmental signals to stimulate or repress the flux from primary to secondary metabolism.

9.5 *A silent partner*

The cognate phosphatase encoded by SCO3845 at first sight appears to be the most likely candidate for the dephosphorylation of the PknB like protein encoded by SCO3848. After all, the overwhelming evidence from study of all the überoperons carrying paired PknB like kinase – phosphatase genes has consistently shown that the phosphatase will dephosphorylate the kinase and that deletion of this phosphatase has severe implications for cell metabolism (Gaidenko, Kim, and Price 2002; Madec et al. 2002; Boitel et al.

2003; Chopra et al. 2003; Rajagopal, Clancy, and Rubens 2003; Echenique et al. 2004). The evidence presented in Chapter 7 of this thesis is therefore surprising, in that removal of SCO3845 seems to present the *S. coelicolor* colony with very few morphological defects, and certainly not the opposite phenotype of retarded aerial development and antibiotic production which logically would be expected. Two possibilities present themselves: Either the phosphatase is not the agent that dephosphorylates *S. coelicolor* PknB, or a different phosphatase encoded in the *S. coelicolor* genome is able to perform this task in place of the protein coded for by SCO3845. The latter solution seems more likely, especially when comparisons are made of the number of phosphatases present in other genomes. *M. tuberculosis* has 1 such phosphatase, *B. subtilis* has 4, whilst *S. coelicolor* has 45 (data obtained from using SCO3845 amino acid sequence in a SMART search for represented domains). This redundancy of function is not a new phenomenon in *S. coelicolor*, and it is probably due to the complex soil environment it is native to. Further to the redundancy argument, bioinformatic analysis of the *S. coelicolor* complement of phosphatases reveals SCO7530 as the most likely candidate for the role of SCO3845 surrogate. This gene codes for a putative member of the MerR family of transcriptional regulators, thus providing a possible link between dephosphorylation of PknB and regulation of transcription.

9.6 *FHA Domains: Key transducers of STPK signalling.*

9.6.1 How do FHA domains work?

The first FHA domains were identified as part of nuclear signalling systems in eukaryotes (Hofmann and Bucher 1995) although unlike STPKs, they were simultaneously

recognised in prokaryotes. The best studied eukaryotic example is the Rad53 – Dun1 checkpoint kinase system in *Saccharomyces cerevisiae*. The system relies on ser/thr/tyr kinases that are physically coupled to FHA domains. The FHA domains have been shown to bind to phosphorylated peptides (Durocher et al. 1999), and presumably modulate the subsequent phosphorylation of targets by the autophosphorylated kinases. Bacterial FHA domains are primarily found in the Gram – positives, and they are found in gene locales that implicate them in a wide variety of cellular processes from carbohydrate storage to sporulation (Pallen, Chaudhuri, and Khan 2002). FHA domains share structural similarities in the form of an 11 strand β – sandwich. Their mode of action relies on the recognition of preferential amino acid sequences in their target proteins (Durocher et al. 2000), incorporating a phosphorylated threonine residue flanked by a specific set of amino acids. Thus far the specific mode of action of FHA domains is unknown, however there is mounting evidence from studies of the *Mycobacteria* that they are phosphorylated by STPKs (Molle et al. 2003; Molle et al. 2004; Curry et al. 2005; Grundner, Gay, and Alber 2005; Villarino et al. 2005; Alderwick et al. 2006). The weight of evidence thus far suggests that PknB can phosphorylate the FHA domains coded by the mycobacterial analogues to SCO3843 and SCO3844. Therefore, there is a high probability that the same mode of action occurs when other conserved überoperons of this type are found in other organisms.

9.6.2 Double mutant phenotype: a metabolic brake removed?

The phenotype of early aerial development and increased antibiotic production found in DSCO3843/44 is at first reminiscent of the situation found in DSCO3848. However, once

the initial similarities are noted, the differences between the two mutants are quite marked.

9.6.2.1 pH control

When compared to the wild type, DSCO3843/44 displays a retarded ability to change the surrounding pH of the media on which it is grown. The M145 parental strain increases the pH of the media on which it is grown, whilst DSCO3843/44 either reduces the pH or does not change it. The explanation for this must lie with the ability of the *Streptomyces* colony to control its environment to maximise the amount of mycelial growth possible before secondary metabolism and sporulation are initiated. The appearance of aerial hyphae is closely linked to an increase in pH (see Chapter 5 and (Susstrunk et al. 1998)), and this developmental sensitivity to changes in pH is lost in DSCO3843/44. This implies a role for the FHA domains in the timing of switches from primary to secondary metabolism, and this is discussed further below.

9.6.2.2 Why do combinations of glucose, phosphate and citrate repress aerial development?

Combinations of glucose with one of citrate or phosphate buffers at 10mM concentration represses the formation of aerial hyphae in DSCO3843/44. This can then be rescued, either by replacing the carbon source with mannitol, or by changing the buffer to a biologically neutral one such as MES (Norman E. Good 1966). This situation is reminiscent of the situation found from the mutation of *citA*, citrate synthase (Viollier et al. 2001). This mutant showed a blockage in aerial development when grown on media containing glucose, and this was ascribed to excess acidogenesis. The use of a less

acidogenic carbon source, mannitol, resulted in restoration of aerial development. This study placed CitA (citrate synthase) in the developmental cascade of the products of the *bld* genes, somewhere above *bldJ*. The similarity in phenotype between the *citA* mutants and DSCO3843/44 does prompt the question as to whether or not the FHA genes are involved in the regulation of this gene. A tantalising piece of evidence is revealed from the proteomic analysis of parental M145 and DSCO3843/44 phosphoproteomes presented in Chapter 8. The phosphoproteome of DSCO3843/44 contains a detectable species of phosphorylated CitA which is not detectable in the M145 phosphoproteome. This implies, at the very least, that removal of the FHA domains results in a change in the phosphorylation of CitA, although at this stage it is impossible to say whether or not this is from a direct interaction between CitA and the FHA domains, or whether it is the result of a downstream effect. What is certain is that removal of these domains has a dramatic impact on the metabolism of *S. coelicolor*, and that this should be investigated further.

9.6.2.3 Feeding experiments

The difference between the two mutant strains is quite marked when it comes to the differences observed when different metabolic intermediates are used. The absence of the FHA domains encoded by SCO3843 and SCO3844 results in highly accelerated antibiotic production, and in most cases accelerated aerial development (with the exception of Malic acid supplemented media on which DSCO3843/44 has a bald phenotype). The question is: are the metabolic enzymes themselves affected by the lack of FHA domains, or has their absence affected a more global regulatory network? The argument for direct control of metabolic enzymes in the TCA cycle is supported by the phenotypes displayed at 48h for DSCO3843/44 when grown on media supplemented with

malic acid and succinate. These bald phenotypes, coupled to the production of red undecylprodigiosin only suggest a blockage in the metabolic pathway, and diversion of carbon flux into undecylprodigiosin production. However, a further 24h incubation shows us that there is no blockage, just a difference in the timing of actinorhodin production, which again far outstrips that of the parental M145 strain. The multiplicity of metabolic pathways (many as yet undiscovered) for the interconversion of metabolic intermediates, as well as the degree of redundancy present in the system (10 homologues of succinate dehydrogenase alone according to BLAST searches of SCODB) stretches the idea of a direct physical link between the FHA domains and the enzymes involved in the TCA cycle. This then leads to the argument that a global regulatory network has been affected by the absence of the FHA domains, and this is consistent with the theory that SCO3848 encodes a sensor kinase involved in such a network. The purpose of this network must be to regulate the flux of metabolites from primary to secondary metabolism, as in the manner of 'overflow' metabolism expanded on by Hodgson (Hodgson 2000). This theory posits that the control of secondary metabolism is not a clear – cut switch between 'primary' and 'secondary' metabolism, but that core processes such as the TCA cycle can have build-ups of intermediates relieved through their conversion into non-essential products such as antibiotics. The implication here is that the various overflow pathways must be under tight regulation – if one such pathway is activated when an intermediate is in short supply, then the core metabolism of the cell could be disastrously affected. This situation may present itself in the laboratory as an intriguing phenotype (excess antibiotic production for example) however in the complex and nutrient – limited soil environment a defect such as this would be lethal. If the potential involvement of CitA mentioned in

the previous Chapter is taken into account, then it is additional evidence that metabolism is partly controlled by p – ser and p – thr interactions.

9.6.3 Hyphal morphology of DSCO3843/44 – through thick and thin.

The substrate hyphae of DSCO3843/44 are curious in that they show a mixture of normal width hyphae and hyphae which are much thicker, coupled to the fact that several of these thicker hyphae often join together to form the 'cord – like' structures seen in Chapter 6. The mystery is all the more puzzling as there does not seem to be any difference between the live / dead ratio of hyphae found in the mutant as found by the BacLight stain, which suggests that the hyphae are at the same stage of programmed cell death. This leads us to the questions: What makes the thicker hyphae join to form cords, and why don't all the hyphae behave in the same way? The answer must lie in the manner in which the cell wall is constructed. Lack of the FHA domains could impair the ability of new hyphae to branch away from their 'parent' hyphae, implying regulation of the D-ala – D-ala carboxypeptidases used to break the peptidoglycan cross-links that would otherwise hold the new hypha on the surface of the parental one (McDonough et al. 2002). This idea is helped by the observation from the AFM images shown in Chapter 6 that suggest that the thicker hyphae may in fact be two thinner hyphae that have not separated. As the FHA null mutant phenotype is again apparent during the first 48h of colony growth, it can be assumed that the FHA domains are usually active during this period. If like SCO3848 they are subsequently downregulated, then the usual later cell wall regulation would apply and hyphal growth would proceed normally, presumably producing the thinner hyphae that branch away from the thicker ones. This hypothesis requires a deal more biochemical work to be proven or disproven, however there is a

precedent for a PBP being phosphorylated on threonine residues, which does coincidentally involve the same cluster. It has been found that PknB of *M. tuberculosis* phosphorylates PBPA (orthologous to SCO3847) on threonine residues (Dasgupta et al. 2006). Therefore, it is not unreasonable to suggest a similar role for FHA domains.

9.7 *Integration of signals – a model for the pknB cluster*

9.7.1 From outside to inside the cell

The start point of the PknB signalling pathway is most likely found outside the cell and the signal then transduced across the cell membrane and then further transmitted to target proteins via PknB, which is the only viable candidate for fulfilling such an extracellular sensory and intracellular signalling role found in this cluster. The most likely class of molecule that PknB senses is unlinked peptidoglycan monomers, and this is achieved via the PASTA domains, which also help localise the protein to the site of their presence in the cell wall (presumably the hyphal tip in *S. coelicolor*) (Pares et al. 1996; Dessen et al. 2001; Yeats, Finn, and Bateman 2002; Davies, Shang, and Bush 2006; Jones and Dyson 2006). The great multiplicity of Gram – positive peptidoglycan variation in the stem peptide (Schleifer and Kandler 1972) naturally leads to the question of which type is sensed by the PknB PASTA domains? If it is assumed that different stages of cell wall development require different types of peptidoglycan (e.g. septal versus elongatory types), then clues to the type of peptidoglycan monomer sensed may be sought from the timing of PknB expression, and whether or not this coincides with a particular developmental stage. The peak of PknB expression in *S. coelicolor* occurs just before a switch from the elongation of aerial hyphae to their forming spores. It may be assumed

that PknB is sensing the availability of peptidoglycan monomers involved in the elongation of these hyphae, and that a change in signalling is induced when they are no longer available (Figure 9.1).

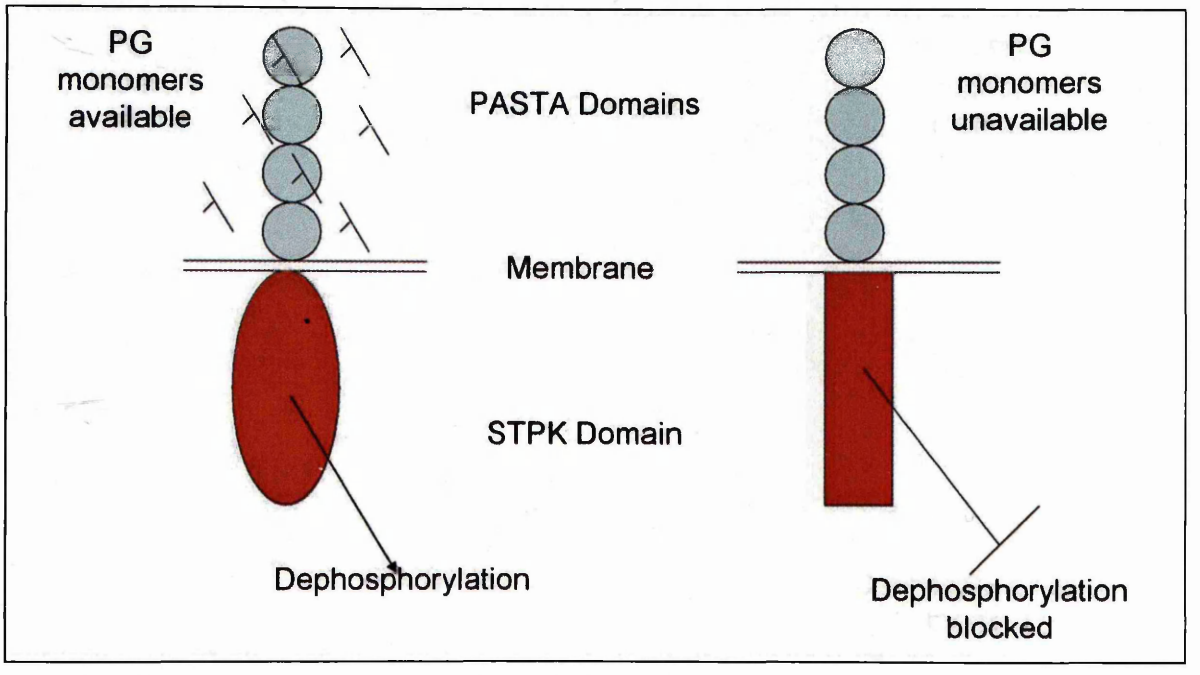


Figure 9.1: A possible model for PknB activity in response to peptidoglycan (PG) monomer availability. In the left hand scenario, hyphal elongation is generating a pool of unlinked peptidoglycan monomers, sensed by the PASTA domains. This results in a conformational change that allows autophosphorylated PknB to be dephosphorylated. In the right hand scenario, the elongation of the hypha has ceased, and the absence of PG monomers causes a conformational change that blocks dephosphorylation.

that PknB is sensing the availability of peptidoglycan monomers involved in the elongation of these hyphae, and that a change in signalling is induced when they are no longer available (Figure 9.1).

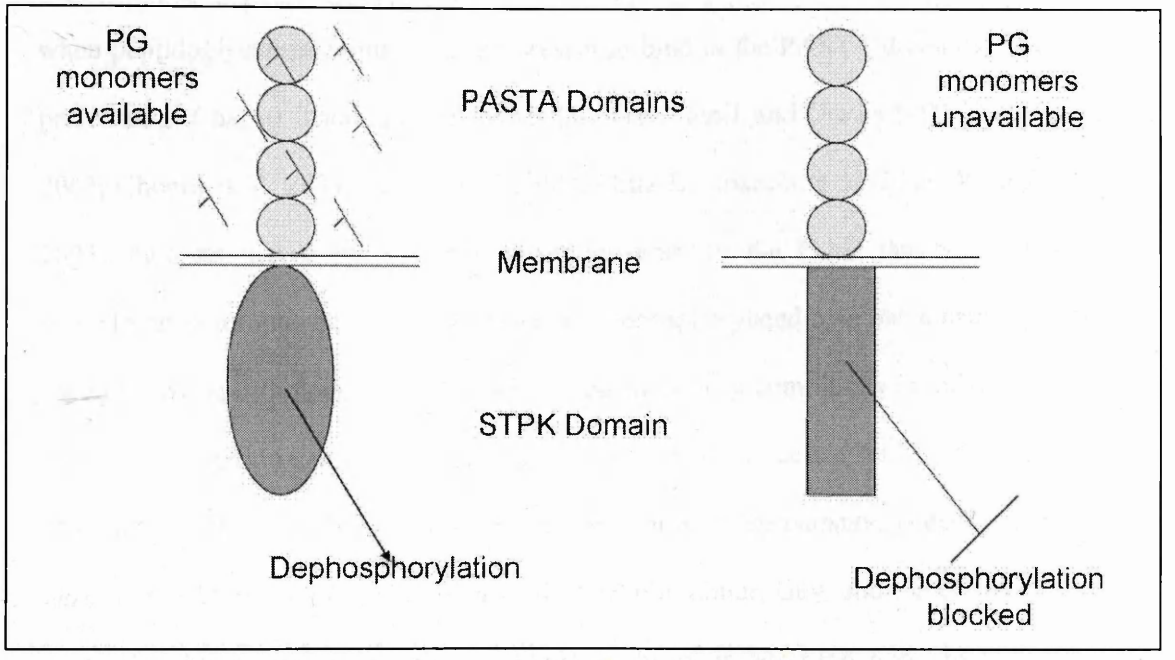


Figure 9.1: A possible model for PknB activity in response to peptidoglycan (PG) monomer availability. In the left hand scenario, hyphal elongation is generating a pool of unlinked peptidoglycan monomers, sensed by the PASTA domains. This results in a conformational change that allows autophosphorylated PknB to be dephosphorylated. In the right hand scenario, the elongation of the hypha has ceased, and the absence of PG monomers causes a conformational change that blocks dephosphorylation.

9.7.2 Dephosphorylation of PknB

The behaviour of purified PknB *in vitro* gives a clue to the 'resting' state of the protein when peptidoglycan monomers are not present to bind to the PASTA domains. The data presented in Chapter 8 and in other studies (Av-Gay, Jamil, and Drews 1999; Boitel et al. 2003; Chopra et al. 2003; Madec et al. 2003; Ortiz-Lombardia et al. 2003; Young et al. 2003) sug gests that there is a mechanism universal to the PknB family of proteins whereby they autophosphorylate before being dephosphorylated by a phosphatase protein coded for by an orthologue of SCO3845. An added layer of complexity is added to by the presence of FHA domains in the same genomic environment (coded for by orthologues of SCO3843 and SCO3844), one of which has been shown to be phosphorylated by PknB *in vitro* by the *M. tuberculosis* orthologue of PknB (Grundner, Gay, and Alber 2005). This study also demonstrated the ability of PknB to phosphorylate FHA domains in proteins encoded beyond the genomic environment of the PknB cluster (those from Rv1747, orthologous to SCO1806). PknB was also shown to phosphorylate an FHA domain carried by GarA, orthologous to SCO6079 (Villarino et al. 2005). This ability of PknB to phosphorylate a wide variety of potential substrates leads to the conclusion that it is a part of a pleiotropic network of STPKs sensing various environmental factors (Av-Gay and Everett 2000) and then responding to them via the phosphorylation of FHA domains, which then go on to modulate the behaviour of further proteins. It is also likely that the mode of action of STPKs is similar, with dimerisation playing an as yet unknown role in the modulation of the transfer of phosphate (Madec et al. 2002; Young et al. 2003; Gay, Ng, and Alber 2006; Wehenkel et al. 2006). It could be that the dimerisation is analogous

to the need for dimerisation of mammalian receptor kinases before the transduction of signals (Heldin 1995), especially when the possibility of a direct evolutionary link between the actinobacterial and eukaryotic signalling machinery is taken into account (Cavalier-Smith 2002).

It has also been shown that PknB has the ability to phosphorylate the PBP coded for by the *M. tuberculosis* orthologue of SCO3847 (Dasgupta et al. 2006). This study showed that a mutated version of this PBP lacking the identified phosphorylation site resulted in a similar phenotype to a strain carrying a null mutation of the PBP. The authors suggested that phosphorylation by the kinase allowed the PBP to participate in cell division. A potential flaw in this reasoning is that the PBP is predicted to be a secreted protein. PknB kinase activity is predicted to be exclusively intra – cellular. Therefore, unless phosphorylation occurs within the cell before secretion, some other method of regulation must control this protein. A possibility is that the SCO3847 PBP is phosphorylated by the FHA domain coded by SCO3844, which is itself predicted to be an extracellular protein. DSCO3848 did not show a difference in cell division (i.e. sporulation) to the M145 parental strain other than in the timing of aerial development, thus showing any phosphorylation of the SCO3847 associated PBP is non – essential. H owever, it is possible that there is a degree of redundancy in *S. coelicolor* cell division as a BLAST search reveals 12 similar proteins to that coded by SCO3847 within the proteome.

9.7.3 Regulatory events after dephosphorylation

Currently, the evidence for the final targets of PknB and its associated FHA domain proteins is not definitive. There is evidence of a profound shift in flux between primary

and secondary metabolism (in favour of secondary metabolism) when the FHA domains are removed from the system, thus a conclusion may be drawn that the FHA domains regulate some aspect of this flux. The proteomic data presented in Chapter 8 has a piece of evidence to support this, with the presence of phosphorylated CitA, citrate synthase, in the phosphoproteome of DSCO3843/44 but not in that of either M145 or DSCO3848. Citrate synthase catalyses the following reaction:

$$\text{Oxaloacetate} + \text{Acetyl CoA} \rightarrow \text{Citrate} + \text{CoA}$$

Therefore, if this enzyme were to be inhibited in its function by phosphorylation, a buildup of Acetyl CoA would occur, and this compound is one of the building blocks of actinorrhodin. A *citA* mutant is unable to produce pigmented antibiotics or aerial mycelium on glucose – containing media (Viollier et al. 2001), however this mutant was able to produce aerial mycelium and pigmented antibiotics when the growth media was buffered to neutrality. The presence of *phosphorylated* CitA in the DSCO3843/44 phosphoproteome suggests that regulation involves maintaining dephosphorylation of this enzyme. The idea of phosphorylation of CitA being used as a method of post-translationary control is supported by the detection of its phosphorylated orthologue in *C. glutamicum* (Bendt et al. 2003). This study also showed that the majority of the phosphoproteome is dedicated to primary metabolism, supporting the idea that STPK – FHA domain mediated phosphosignalling can have a dramatic impact on the flux of metabolism.

9.7.3.1 Further phosphoproteomic differences

The phosphorylated proteins detected that were common to M145 and both mutant strains studied showed a difference in the regulation of RNA polymerase in the mutants, with

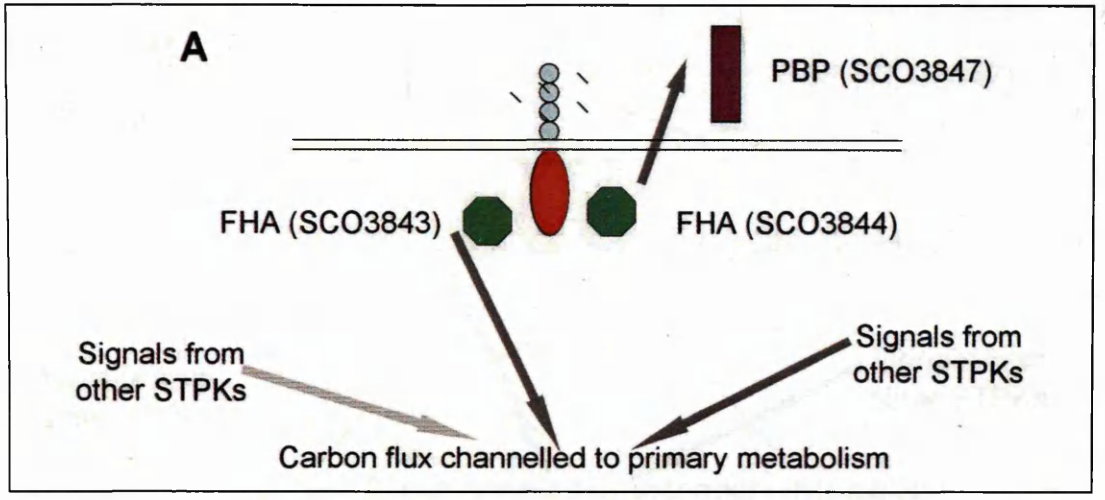
more abundant peptides being detected in the M145 parental strain. This shows that more phosphorylated RpoC was present in M145. This either reflects a higher level of transcription, or a difference in its regulation by binding of ppGpp (Hesketh, Sun, and Bibb 2001; Xu et al. 2002), suggesting a common alteration in RNA synthesis between the two mutant strains, most likely due to the early onset of secondary metabolism, and this is supported by the presence of phosphorylated RpoA in both mutant strains but not in M145. By the same token, DSCO3843/44 contains more phosphorylated CobN, cobalamin biosynthesis protein than the other two strains which indicates a difference in the synthesis of coenzyme B12. A difference in coenzyme levels would indicate a different pattern of enzyme activity in the mutant strain, and this is likely to be reflected in the more vigorous secondary metabolism of this mutant.

Phosphorylated proteins shared between the two mutant strains show no differences in abundance, and their predicted functions show that primary metabolism, RNA synthesis and electron transport are differently regulated in the mutant strains than in M145.
The two phosphorylated proteins found common to DSCO3848 and M145 exclusively show that the putative ATPase and GpsI are not regulated by PknB, and their absence from DSCO3843/44 suggests the possibility they are involved in FHA domain mediated interactions.

The most interesting phosphorylated protein found commonly between M145 and DSCO3843/44 is MurA, UDP-N-acetylglucosamine transferase, involved in the

manufacture of the cell wall. Again, the absence of this phosphoprotein from DSCO3848 allows the possibility of its regulation by PknB.

Of the proteins found exclusively in the phosphoproteomes of each strain, it is apparent that the majority have not been characterized other than through sequence comparison. The functions of the proteins identified are involved mainly in primary metabolism and protein synthesis, and the presence of a putative Spo0M orthologue in the DSCO3848 phosphoproteome again shows a possible link with cell division.



Figuure 9.2a: Putative scenario in which PknB acts as part of a pleiotropic STPK network regulating aspects of metabolism and possibly cell division. The active form of the kinase autophosphorylates and is dephosphorylated in the presence of PG monomers. The dephosphorylation is probably aided by one or more PP2C type phosphatases (not shown). The phosphate group is then used in a signalling cascade that channels metabolic intermediates towards primary metabolism to provide enough energy and precursors for hyphal extension. Also shown is the possibility of regulation of the secreted SCO3847 PBP by the secreted SCO3844 FHA domain.

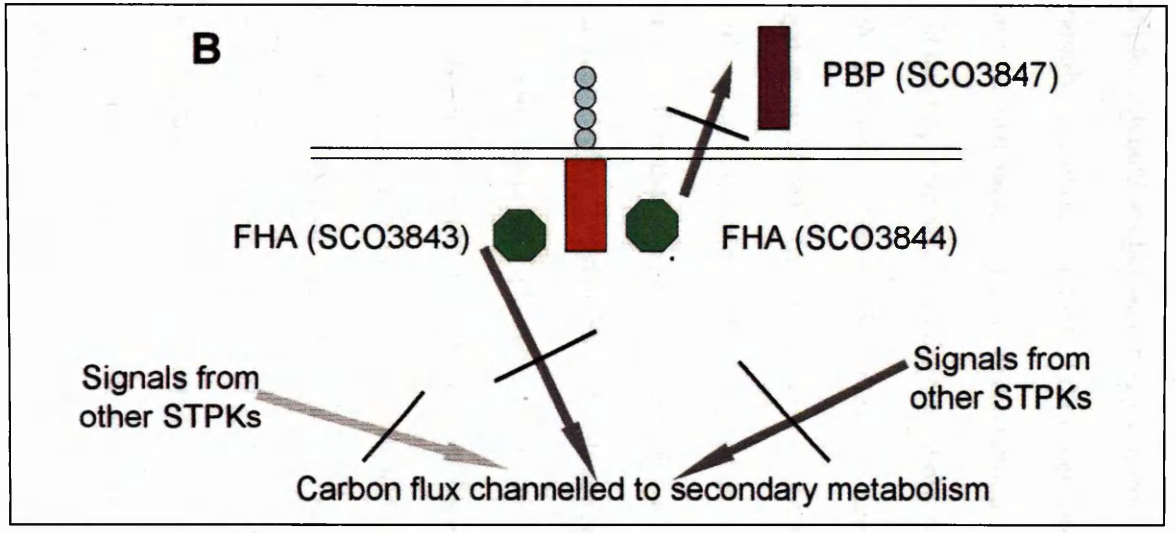


Figure 9.2b: When PG monomers become less abundant due to the cessation of hyphal extension, the STPK domain is no longer dephosphorylated. The FHA domains are therefore no longer used to channel carbon flux towards primary metabolism, and this is part of the regulatory framework that biases metabolism towards secondary metabolism. The cessation of regulation of the PBP would make sense if it were involved in hyphal extension only and not septum development.

9.7.4 Pleiotropic networks involving STPKs

The idea that actinobacterial STPKs are part of a global regulatory network involving complex interactions between many different proteins phosphorylated on ser / thr motifs is rapidly emerging. The first of these networks was found in *S. coelicolor*, whereby the transcriptional factor AfsR is phosphorylated by a number of STPKs and this controls aerial development and secondary metabolism (Matsumoto et al. 1994; Floriano and Bibb 1996; Umeyama et al. 1999; Umeyama, Lee, and Horinouchi 2002; Horinouchi 2003; Sawai et al. 2004). Another network discovered recently in *M. tuberculosis* implicates STPKs in mycolic acid synthesis (Molle et al. 2006), and this study also showed for the first time that the mycobacterial orthologue of the phosphatase encoded by SCO3845 (PstP) can dephosphorylate PknB substrates (KasA and KasB). This study also showed a large amount of cross – phosphorylation between STPKs and their substrates, with each substrate able to be phosphorylated by multiple STPKs. The promiscuous nature of STPK phosphorylation does not aid a neat categorisation of PknB, indeed it seems that the only sure fact about this protein is its ability to phosphorylate a wide variety of substrates. The most logical conclusion must therefore be that in *S. coelicolor* the PknB orthologue translated from SCO3848 is part of a wide network of proteins which control the metabolism of the bacterial cell to facilitate optimum growth at various developmental stages of the life cycle.

It is becoming apparent that current tools for analysing protein interactions are too focussed on the single interactions between two protein partners. The development of

increasingly powerful proteomic methods will lead to a much better understanding of the post – translationary regulation of proteins.

9.8 *Conclusion*

The signalling genes of the p*knB* gene cluster are conserved across many species of bacteria belonging to the Actinobacteria and Firmicutes, and this conservation of gene order implies a conserved function. The key to the cluster is the STPK PknB, a transmembrane kinase with a number of extracellular PASTA domains and an intracellular STPK domain. The PASTA domains are thought to sense unlinked peptidoglycan monomers, although whether or not this is a device to locate the kinase to the site of cell wall synthesis (as PASTA domains are thought to do in PBP2X) or purely as a signalling trigger requires more study. In *S. coelicolor* the orthologue of PknB, coded for by SCO3848, is one of two such kinases present in the genome (the other being SCO2110). This leads to the conclusion that there is a certain redundancy of function, especially when a *pknB* mutant in the genetically similar *S. avermitilis* did not show a detectable phenotype. The effects of disrupting SCO3848 were those of early onset of aerial development and secondary metabolism, and this phenotype was similar to that displayed by the disruption of SCO3843 and SCO3844. These two genes encode FHA domain containing proteins, which mediate p-ser and p-thr interactions. The absence of these FHA domains led to a dramatic increase in the amount of secondary metabolites produced, leading to speculation that phosphorylation plays an important role in the regulation of metabolic flux. This was supported by identification of a large number of metabolism – related genes in the phosphoproteomes of M145 and mutant strains, with

differences between M145 and mutants being exclusively related to metabolic activity and housekeeping functions. When studies in *Mycobacteria* are taken into account the emerging (and daunting!) picture is one of a richly complex network of phosphorylated proteins sensing environmental signals using sensor STPKs and then transducing that signal via FHA domains and PstP type phosphatases to specific targets. These networks also coalesce at specific nodal proteins such as AfsR in *S. coelicolor*. Thus subtle environmental cues can be met with an equally subtle responses from the cell. In the case of SCO3848, its kinase is expressed until the onset of sporulation, implicating it in the maintenance of cellular conditions conducive to the elongation of the hyphae. Upon the onset of sporulation, PknB signals a shift in the metabolic flux of the cell in favour of secondary metabolism.

9.9 *Further work*

The most pressing question is to definitively answer whether or not CitA is regulated by the FHA domains, and this could be investigated using a combination of analysis of recombinant protein interactions using surface plasmon resonance for example, as well as bioinformatic identification of potential phosphorylation sites and the subsequent production of site – specific mutant strains of *S. coelicolor*. Surface plasmon resonance experiments could also be used on PASTA domains to identify the moieties sensed by these domains. Microarray analysis of the mutant strains would also yield useful data on the genes affected by the absence of the STPK and FHA genes, and would also help the identification of new targets.

9.10 *References for Chapter 9*

Alderwick, L. J., V. Molle, L. Kremer, A. J. Cozzone, T. R. Dafforn, G. S. Besra, and K. Futterer. 2006. Molecular structure of EmbR, a response element of Ser/Thr kinase signaling in Mycobacterium tuberculosis
10.1073/pnas.0507766103. PNAS **103**:2558-2563.

Av-Gay, Y., and M. Everett. 2000. The eukaryotic-like Ser/Thr protein kinases of Mycobacterium tuberculosis. Trends Microbiol **8**:238-244.

Av-Gay, Y., S. Jamil, and S. J. Drews. 1999. Expression and characterization of the Mycobacterium tuberculosis serine/threonine protein kinase PknB. Infect Immun **67**:5676-5682.

Bendt, A. K., A. Burkovski, S. Schaffer, M. Bott, M. Farwick, and T. Hermann. 2003. Towards a phosphoproteome map of Corynebacterium glutamicum. PROTEOMICS **3**:1637-1646.

Bentley, S. D., K. F. Chater, A. M. Cerdeno-Tarraga, G. L. Challis, N. R. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C. W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C. H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Neil, E. Rabbinowitsch, M. A. Rajandream, K. Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B. G. Barrell, J. Parkhill, and D. A. Hopwood. 2002. Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). Nature **417**:141-147.

Boitel, B., M. Ortiz-Lombardia, R. Duran, F. Pompeo, S. T. Cole, C. Cervenansky, and P. M. Alzari. 2003. PknB kinase activity is regulated by phosphorylation in two Thr residues and dephosphorylation by PstP, the cognate phospho-Ser/Thr phosphatase, in Mycobacterium tuberculosis. Mol Microbiol **49**:1493-1508.

Boyle, D. S., M. M. Khattar, S. G. Addinall, J. Lutkenhaus, and W. D. Donachie. 1997. ftsW is an essential cell-division gene in Escherichia coli. Mol Microbiol **24**:1263-1273.

Cavalier-Smith, T. 2002. The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. Int J Syst Evol Microbiol **52**:297-354.

Chen, C. W., C. H. Huang, H. H. Lee, H. H. Tsai, and R. Kirby. 2002. Once the circle has been broken: dynamics and evolution of Streptomyces chromosomes. Trends Genet **18**:522-529.

Chopra, P., B. Singh, R. Singh, R. Vohra, A. Koul, L. S. Meena, H. Koduri, M. Ghildiyal, P. Deol, T. K. Das, A. K. Tyagi, and Y. Singh. 2003. Phosphoprotein phosphatase of Mycobacterium tuberculosis dephosphorylates serine-threonine kinases PknA and PknB. Biochem Biophys Res Commun **311**:112-120.

Curry, J. M., R. Whalan, D. M. Hunt, K. Gohil, M. Strom, L. Rickman, M. J. Colston, S. J. Smerdon, and R. S. Buxton. 2005. An ABC Transporter Containing a Forkhead-Associated Domain Interacts with a Serine-Threonine Protein Kinase and Is Required for Growth of Mycobacterium tuberculosis in Mice. Infect. Immun. **73**:4471-4477.

Da Lage, J. L., G. Feller, and S. Janecek. 2004. Horizontal gene transfer from Eukarya to bacteria and domain shuffling: the alpha-amylase model. Cell Mol Life Sci **61**:97-109.

Dasgupta, A., P. Datta, M. Kundu, and J. Basu. 2006. The serine/threonine kinase PknB of Mycobacterium tuberculosis phosphorylates PBPA, a penicillin-binding protein required for cell division
10.1099/mic.0.28630-0. Microbiology **152**:493-504.

Davies, T. A., W. Shang, and K. Bush. 2006. Activities of Ceftobiprole and Other {beta}-Lactams against Streptococcus pneumoniae Clinical Isolates from the United States with Defined Substitutions in Penicillin-Binding Proteins PBP 1a, PBP 2b, and PBP 2x
10.1128/AAC.00238-06. Antimicrob. Agents Chemother. **50**:2530-2532.

Del Sol, R., J. G. Mullins, N. Grantcharova, K. Flardh, and P. Dyson. 2006. Influence of CrgA on Assembly of the Cell Division Protein FtsZ during Development of Streptomyces coelicolor. J Bacteriol **188**:1540-1550.

Del Sol, R., A. Pitman, P. Herron, and P. Dyson. 2003. The product of a developmental gene, crgA, that coordinates reproductive growth in Streptomyces belongs to a novel family of small actinomycete-specific proteins. J Bacteriol **185**:6678-6685.

Dessen, A., N. Mouz, E. Gordon, J. Hopkins, and O. Dideberg. 2001. Crystal Structure of PBP2x from a Highly Penicillin-resistant Streptococcus pneumoniae Clinical Isolate. A MOSAIC FRAMEWORK CONTAINING 83 MUTATIONS
10.1074/jbc.M107608200. J. Biol. Chem. **276**:45106-45112.

Doull, J. L., and L. C. Vining. 1990. Nutritional control of actinorhodin production byStreptomyces coelicolor A3(2): suppressive effects of nitrogen and phosphate. Applied Microbiology and Biotechnology **V32**:449-454.

Durocher, D., J. Henckel, A. R. Fersht, and S. P. Jackson. 1999. The FHA domain is a modular phosphopeptide recognition motif. Mol Cell **4**:387-394.

Durocher, D., I. A. Taylor, D. Sarbassova, L. F. Haire, S. L. Westcott, S. P. Jackson, S. J. Smerdon, and M. B. Yaffe. 2000. The molecular basis of FHA domain:phosphopeptide binding specificity and implications for phospho-dependent signaling mechanisms. Mol Cell **6**:1169-1182.

Echenique, J., A. Kadioglu, S. Romao, P. W. Andrew, and M.-C. Trombe. 2004. Protein Serine/Threonine Kinase StkP Positively Controls Virulence and Competence in Streptococcus pneumoniae. Infect. Immun. **72**:2434-2437.

Fernandez, P., B. Saint-Joanis, N. Barilone, M. Jackson, B. Gicquel, S. T. Cole, and P. M. Alzari. 2006. The Ser/Thr Protein Kinase PknB Is Essential for Sustaining Mycobacterial Growth. J. Bacteriol. **188**:7778-7784.

Flardh, K. 2003. Essential role of DivIVA in polar growth and morphogenesis in Streptomyces coelicolor A3(2). Molecular Microbiology **49**:1523-1536.

Floriano, B., and M. Bibb. 1996. afsR is a pleiotropic but conditionally required regulatory gene for antibiotic production in Streptomyces coelicolor A3(2)
doi:10.1046/j.1365-2958.1996.6491364.x. Molecular Microbiology **21**:385-396.

Gaidenko, T. A., T. J. Kim, and C. W. Price. 2002. The PrpC serine-threonine phosphatase and PrkC kinase have opposing physiological roles in stationary-phase Bacillus subtilis cells. J Bacteriol **184**:6109-6114.

Garcia-Bustos, J., and A. Tomasz. 1990. A Biological Price of Antibiotic Resistance: Major Changes in the Peptidoglycan Structure of Penicillin-Resistant Pneumococci
10.1073/pnas.87.14.5415. PNAS **87**:5415-5419.
Gay, L. M., H.-L. Ng, and T. Alber. 2006. A Conserved Dimer and Global Conformational Changes in the Structure of apo-PknE Ser/Thr Protein Kinase from Mycobacterium tuberculosis. Journal of Molecular Biology **360**:409-420.
Granger, D., G. Boily-Larouche, P. Turgeon, K. Weiss, and M. Roger. 2005. Genetic analysis of pbp2x in clinical Streptococcus pneumoniae isolates in Quebec, Canada. J. Antimicrob. Chemother. **55**:832-839.
Grundner, C., L. M. Gay, and T. Alber. 2005. Mycobacterium tuberculosis serine/threonine kinases PknB, PknD, PknE, and PknF phosphorylate multiple FHA domains. Protein Sci **14**:1918-1921.
Heldin, C.-H. 1995. Dimerization of cell surface receptors in signal transduction. Cell **80**:213-223.
Hesketh, A., J. Sun, and M. Bibb. 2001. Induction of ppGpp synthesis in Streptomyces coelicolor A3(2) grown under conditions of nutritional sufficiency elicits actII-ORF4 transcription and actinorhodin biosynthesis. Molecular Microbiology **39**:136-144.
Hodgson, D. A. 2000. Primary metabolism and its control in streptomycetes: a most unusual group of bacteria. Adv Microb Physiol **42**:47-238.
Hofmann, K., and P. Bucher. 1995. The FHA domain: a putative nuclear signalling domain found in protein kinases and transcription factors. Trends Biochem Sci **20**:347-349.
Horinouchi, S. 2003. AfsR as an integrator of signals that are sensed by multiple serine/threonine kinases in Streptomyces coelicolor A3(2). J Ind Microbiol Biotechnol **30**:462-467.
Huynen, M., B. Snel, I. Lathe, Warren, and P. Bork. 2000. Exploitation of gene context. Current Opinion in Structural Biology **10**:366-370.
Ikeda, H., J. Ishikawa, A. Hanamoto, M. Shinose, H. Kikuchi, T. Shiba, Y. Sakaki, M. Hattori, and S. Omura. 2003. Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis. Nat Biotechnol **21**:526-531.
Jones, G., and P. Dyson. 2006. Evolution of Transmembrane Protein Kinases Implicated in Coordinating Remodeling of Gram-Positive Peptidoglycan: Inside versus Outside. J. Bacteriol. **188**:7470-7476.
Kang, C.-M., D. W. Abbott, S. T. Park, C. C. Dascher, L. C. Cantley, and R. N. Husson. 2005. The Mycobacterium tuberculosis serine/threonine kinases PknA and PknB: substrate identification and regulation of cell shape. Genes Dev.:gad.1311105.
Khattar, M. M., S. G. Addinall, K. H. Stedul, D. S. Boyle, J. Lutkenhaus, and W. D. Donachie. 1997. Two polypeptide products of the Escherichia coli cell division gene ftsW and a possible role for FtsW in FtsZ function. J Bacteriol **179**:784-793.
Kodani, S., M. E. Hudson, M. C. Durrant, M. J. Buttner, J. R. Nodwell, and J. M. Willey. 2004. From The Cover: The SapB morphogen is a lantibiotic-like peptide derived from the product of the developmental gene ramS in Streptomyces coelicolor
10.1073/pnas.0404220101. PNAS **101**:11448-11453.

Kurland, C. G., B. Canback, and O. G. Berg. 2003. Horizontal gene transfer: a critical view. Proc Natl Acad Sci U S A **100**:9658-9662.
Lathe, W. C., 3rd, B. Snel, and P. Bork. 2000. Gene context conservation of a higher order than operons. Trends Biochem Sci **25**:474-479.
Madec, E., A. Laszkiewicz, A. Iwanicki, M. Obuchowski, and S. Seror. 2002. Characterization of a membrane-linked Ser/Thr protein kinase in Bacillus subtilis, implicated in developmental processes. Mol Microbiol **46**:571-586.
Madec, E., A. Stensballe, S. Kjellstrom, L. Cladiere, M. Obuchowski, O. N. Jensen, and S. J. Seror. 2003. Mass Spectrometry and Site-directed Mutagenesis Identify Several Autophosphorylated Residues Required for the Activity of PrkC, a Ser/Thr Kinase from Bacillus subtilis. Journal of Molecular Biology **330**:459-472.
Matsumoto, A., S. K. Hong, H. Ishizuka, S. Horinouchi, and T. Beppu. 1994. Phosphorylation of the AfsR protein involved in secondary metabolism in Streptomyces species by a eukaryotic-type protein kinase. Gene **146**:47-56.
McDonough, M. A., J. W. Anderson, N. R. Silvaggi, R. F. Pratt, J. R. Knox, and J. A. Kelly. 2002. Structures of two kinetic intermediates reveal species specificity of penicillin-binding proteins. J Mol Biol **322**:111-122.
Molle, V., A. K. Brown, G. S. Besra, A. J. Cozzone, and L. Kremer. 2006. The Condensing Activities of the Mycobacterium tuberculosis Type II Fatty Acid Synthase Are Differentially Regulated by Phosphorylation
10.1074/jbc.M601691200. J. Biol. Chem. **281**:30094-30103.
Molle, V., L. Kremer, C. Girard-Blanc, G. S. Besra, A. J. Cozzone, and J. F. Prost. 2003. An FHA phosphoprotein recognition domain mediates protein EmbR phosphorylation by PknH, a Ser/Thr protein kinase from Mycobacterium tuberculosis. Biochemistry **42**:15300-15309.
Molle, V., D. Soulat, J. M. Jault, C. Grangeasse, A. J. Cozzone, and J. F. Prost. 2004. Two FHA domains on an ABC transporter, Rv1747, mediate its phosphorylation by PknF, a Ser/Thr protein kinase from Mycobacterium tuberculosis. FEMS Microbiol Lett **234**:215-223.
Morlot, C., M. Noirclerc-Savoye, A. Zapun, O. Dideberg, and T. Vernet. 2004. The d,d-carboxypeptidase PBP3 organizes the division process of Streptococcus pneumoniae. Molecular Microbiology **51**:1641-1648.
Morlot, C., A. Zapun, O. Dideberg, and T. Vernet. 2003. Growth and division of Streptococcus pneumoniae: localization of the high molecular weight penicillin-binding proteins during the cell cycle. Molecular Microbiology **50**:845-855.
Nguyen, K. T., J. M. Willey, L. D. Nguyen, L. T. Nguyen, P. H. Viollier, and C. J. Thompson. 2002. A central regulator of morphological differentiation in the multicellular bacterium Streptomyces coelicolor. Mol Microbiol **46**:1223-1238.
Norman E. Good, G. D. W., Wilhelmina Winter, Thomas N. Connolly, Seikichi Izawa, Raizada M. M. Singh. 1966. Hydrogen Ion Buffers for Biological Research. Biochemistry **5**:467 - 477.
Ortiz-Lombardia, M., F. Pompeo, B. Boitel, and P. M. Alzari. 2003. Crystal structure of the catalytic domain of the PknB serine/threonine kinase from Mycobacterium tuberculosis. J Biol Chem **278**:13094-13100.
Pallen, M., R. Chaudhuri, and A. Khan. 2002. Bacterial FHA domains: neglected players in the phospho-threonine signalling game? Trends Microbiol **10**:556-563.

Paolo, S. S., J. Huang, S. N. Cohen, and C. J. Thompson. 2006. rag genes: novel components of the RamR regulon that trigger morphological differentiation in Streptomyces coelicolor
doi:10.1111/j.1365-2958.2006.05304.x. Molecular Microbiology **61**:1167-1186.
Pares, S., N. Mouz, Y. Petillot, R. Hakenbeck, and O. Dideberg. 1996. X-ray structure of Streptococcus pneumoniae PBP2x, a primary penicillin target enzyme. **3**:284-289.
Pastoret, S., C. Fraipont, T. den Blaauwen, B. Wolf, M. E. G. Aarsman, A. Piette, A. Thomas, R. Brasseur, and M. Nguyen-Disteche. 2004. Functional Analysis of the Cell Division Protein FtsW of Escherichia coli. J. Bacteriol. **186**:8370-8379.
Rajagopal, L., A. Clancy, and C. E. Rubens. 2003. A Eukaryotic Type Serine/Threonine Kinase and Phosphatase in Streptococcus agalactiae Reversibly Phosphorylate an Inorganic Pyrophosphatase and Affect Growth, Cell Segregation, and Virulence
10.1074/jbc.M212747200. J. Biol. Chem. **278**:14429-14441.
Rogozin, I. B., K. S. Makarova, J. Murvai, E. Czabarka, Y. I. Wolf, R. L. Tatusov, L. A. Szekely, and E. V. Koonin. 2002. Connected gene neighborhoods in prokaryotic genomes. Nucl. Acids Res. **30**:2212-2223.
Sassetti, C. M., D. H. Boyd, and E. J. Rubin. 2003. Genes required for mycobacterial growth defined by high density mutagenesis. Mol Microbiol **48**:77-84.
Sawai, R., A. Suzuki, Y. Takano, P. C. Lee, and S. Horinouchi. 2004. Phosphorylation of AfsR by multiple serine/threonine kinases in Streptomyces coelicolor A3(2). Gene **334**:53-61.
Schleifer, K. H., and O. Kandler. 1972. Peptidoglycan types of bacterial cell walls and their taxonomic implications. Bacteriol Rev **36**:407-477.
Sola-Landa, A., R. S. Moura, and J. F. Martin. 2003. The two-component PhoR-PhoP system controls both primary metabolism and secondary metabolite biosynthesis in Streptomyces lividans
10.1073/pnas.0931429100. PNAS **100**:6133-6138.
Susstrunk, U., J. Pidoux, S. Taubert, A. Ullmann, and C. J. Thompson. 1998. Pleiotropic effects of cAMP on germination, antibiotic biosynthesis and morphological development in Streptomyces coelicolor. Molecular Microbiology **30**:33-46.
Umeyama, T., P. C. Lee, and S. Horinouchi. 2002. Protein serine/threonine kinases in signal transduction for secondary metabolism and morphogenesis in Streptomyces. Appl Microbiol Biotechnol **59**:419-425.
Umeyama, T., P. C. Lee, K. Ueda, and S. Horinouchi. 1999. An AfsK/AfsR system involved in the response of aerial mycelium formation to glucose in Streptomyces griseus. Microbiology **145 (Pt 9)**:2281-2292.
Villarino, A., R. Duran, A. Wehenkel, P. Fernandez, P. England, P. Brodin, S. T. Cole, U. Zimny-Arndt, P. R. Jungblut, C. Cervenansky, and P. M. Alzari. 2005. Proteomic Identification of M. tuberculosis Protein Kinase Substrates: PknB Recruits GarA, a FHA Domain-containing Protein, Through Activation Loop-mediated Interactions. Journal of Molecular Biology **350**:953-963.
Viollier, P. H., W. Minas, G. E. Dale, M. Folcher, and C. J. Thompson. 2001. Role of acid metabolism in Streptomyces coelicolor morphological differentiation and antibiotic biosynthesis. J Bacteriol **183**:3184-3192.
Wehenkel, Fernandez, Bellinzoni, Catherinot, Barilone, Labesse, Jackson, and Alzari. 2006. The structure of PknB in complex with mitoxantrone, an ATP-competitive

inhibitor, suggests a mode of protein kinase regulation in mycobacteria. FEBS Letters **580**:3018-3022.

Wolf, Y. I., I. B. Rogozin, A. S. Kondrashov, and E. V. Koonin. 2001. Genome alignment, evolution of prokaryotic genome organization, and prediction of gene function using genomic context. Genome Res **11**:356-372.

Xu, J. Xu, Tozawa, Y. Tozawa, Lai, C. Lai, Hayashi, H. Hayashi, Ochi, and K. Ochi. 2002. A rifampicin resistance mutation in the rpoB gene confers ppGpp-independent antibiotic production in Streptomyces coelicolor A3(2). Molecular Genetics and Genomics **268**:179-189.

Yeats, C., R. D. Finn, and A. Bateman. 2002. The PASTA domain: a beta-lactam-binding domain. Trends Biochem Sci **27**:438.

Young, T. A., B. Delagoutte, J. A. Endrizzi, A. M. Falick, and T. Alber. 2003. Structure of Mycobacterium tuberculosis PknB supports a universal activation mechanism for Ser/Thr protein kinases. Nat Struct Biol **10**:168-174.

Zheng, Y., R. J. Roberts, and S. Kasif. 2002. Genomic functional annotation using co-evolution profiles of gene clusters. Genome Biol **3**:RESEARCH0060.